# THIS WEEK

EDITORIALS

**WORLD VIEW** Impact factors turn Chinese chemistry into a game of bluff **p.253**

**RODINIA** Antarctica and Texas unlikely neighbours in ancient supercontinent **p.254**

**BREEDERS' CUP** European paper wasps work to inherit starter homes **p.255**



## Start small, think big

*The United Kingdom and others must not overlook the potential for nanotechnology to boost regenerative medicine.*

Andemariam Teklesenbet Beyene has benefited from a multidisciplinary approach to surgery. In June, the postgraduate student's cancer-riddled windpipe was replaced with one that had been created in a laboratory. Scientists at University College London created a glass replica of his trachea and two main bronchi, onto which they deposited layers of a porous polymer nanocomposite. Colleagues in Sweden soaked the structure in a solution of Beyene's bone-marrow stem cells to create a fully synthetic organ for transplantation. Windpipes have been replaced before, but Beyene's operation was unique in that no donor was required and there is no risk of rejection. The procedure was a triumph of cooperation, both between international researchers and between the physical and biological sciences.

Materials science has long been recognized as important in such advances. But the field's role is not confined to building biomimetic scaffolds for tissues and organs. The elasticity and surface topography of substrates can control stem-cell growth and differentiation, so materials research can help scientists to prepare cells for clinical use. And it can unravel the biological mechanisms that direct stem-cell behaviour.

Nanotechnology allows researchers to create materials with exquisitely fine structural detail, which are set to be increasingly useful in biology and medicine. Where surface structure and chemistry can be engineered to the cellular and subcellular level, scientists have unprecedented control in probing the responses of cells to their environment.

A paper in this month's *Nature Materials* reports a nanostructured substrate that can maintain stem-cell viability and allow cells to grow for eight weeks (R. J. McMurray *et al. Nature Mater.* **10,** 637–644; 2011). Cells cultured on these surfaces provide insight into the biomolecular factors that control cell-signalling pathways.

But high-profile success of clinical work using tissue scaffolds could draw attention, and funding, away from such fundamental contributions. This may already be happening, at least in Britain. In July, just over a week after news broke of Beyene's operation, the UK government's Office of Life Sciences published a report called *Taking Stock of Regenerative Medicine in the United Kingdom*. It recognizes that research in the physical sciences is needed to move regenerative medicine from the laboratory to the clinic, and notes that new materials, diagnostics and imaging are needed. But it frames useful materials research solely in terms of engineering scaffolds for delivery and application.

In doing so, the government misses an opportunity to outline a remit to investigate materials and the way they control stem cells. Such an approach would help biomedical engineers to design scaffolds and matrices with cell behaviour as a priority, rather than an afterthought.

And although the paper in *Nature Materials* and the engineered trachea (both UK-based research efforts) show that nanotechnology is important to all areas of regenerative medicine, the word nanotechnology does not feature once in the report's 58 pages.

The United Kingdom has no ongoing funding programme that specifically focuses on nanotechnology. In July, Cientifica, a nanotechnology consultancy based in London, released an assessment of global spending in the field that placed Britain almost last, just above India. Richard Jones, a former nanotechnology adviser to the UK Engineering and Physical Sciences Research Council, argues on his blog that Britain has given up on nanotechnology (go.nature.com/xxoy9x).

To help to reverse the trend, scientists, universities and funders should highlight research areas in which nanotechnologists can contribute to scientific success. Regenerative medicine offers such an opportunity. ■

## Keep it clean

*Integrity guidelines are a good start, but they must be clear and appropriately enforced.*

Maintaining integrity in science seems to be in fashion. US officials are preparing to review policies on scientific integrity drafted by some 19 government agencies. Canada is overhauling its research-integrity policy, and the UK Research Integrity Office is looking for new sources of support after losing government funding. The European Commission has also started to discuss how to include scientific integrity in its research framework agreements.

The US policies will apply not only to agency scientists, but also to political appointees, managers and public-affairs officials. This broad scope is deliberate and intended to avoid a repeat of the alleged abuses of science that occurred under the administration of former US President George W. Bush (see page 262).

Still, to frame such policies so broadly does bring potential trouble — how should officials apply them to scientists? Vague requirements, such as to approach research objectively and to welcome constructive criticism, are good professional codes of conduct — but they must not be interpreted as a stick with which politically conscious managers can beat scientists to suppress inconvenient scientific findings.

It would be too easy for officials to violate the intended spirit of a loosely worded scientific-integrity policy by claiming that scientists had violated the letter of it. The United States classifies scientific misconduct as falsification, fabrication or plagiarism, and the broad concept of integrity must be coordinated carefully with that narrow definition. And, of course, good policies are not enough if they are not implemented wisely. To that end, US federal agencies must continue to improve their practices through training for managers of scientific staff, and ensure that they turn to the proper scientific expertise when misconduct is alleged. ■



© 2011 Macmillan Publishers Limited. All rights reserved



# China's chemists should avoid the Vanity Fair

*Problems in Chinese science are well documented. The slow pace of reform affects chemists more than most, says* **Nai-Xing Wang**.

Change comes slowly in China and necessary change to the way Chinese science works is coming very slowly indeed. Despite suggestions earlier this year that reform was in the air, Chinese administrators still stick to their system of allotting funding according to the impact factors of the journals in which we publish our research. And chemistry is the field most adversely affected by this.

I have worked as a researcher in organic chemistry for many years, and am greatly concerned about the way Chinese chemistry is heading. Chemistry has made a great contribution to human society, but right now it seems that chemistry research in China has turned into a Vanity Fair.

The biggest problem remains the obsession with journal impact factors. Generally speaking, articles in journals with high impact factors are judged to appeal most to readers, but not every paper published in a high-impact-factor journal is high quality, and papers published in lower-ranked journals are never worthless. Yet some administrators in China take a very crude approach: high-impact-factor publications mean excellent work.

Research proposals are judged according to the impact factor of a scientist's previous publications. (And referees are usually selected on these criteria too.) Worse, the salaries of my chemistry colleagues go up or down depending on a complex mathematical formula based on the impact factor of the journals in which we publish our work — which we must supply in a detailed list.

An impact factor is a measure only of the frequency with which the average article in a journal is cited in a given period of time. So it is no surprise to find journals judged to have high impact in fashionable research areas, such as nanotechnology. For example, the impact factor of *Nanotechnology Letters* was 9.99 in 2010, but we should not judge this journal twice as valuable as the *Journal of Organic Chemistry*, with an impact factor of 4.21 in 2010. And some researchers in nanotechnology have not used mainstream techniques such as nuclear magnetic resonance, mass spectrography, infrared or ultraviolet spectroscopy, or X-ray diffraction. They sometimes get by with just one instrument: an electron microscope. Some research on 'novel' materials is nothing more than pictures taken with such a microscope. In contrast, papers in organic synthetic chemistry often feature new synthetic routes and skills, but because some of the reactions involved are common, they are too often dismissed as lacking innovation and novelty.

If a high impact factor is the only goal of chemistry research, then chemistry is no longer science. It is changed to a field of fame and game. There are other effects too. Administrators in almost every university and research institute like to evaluate researchers by their papers at the end of each year. As a result, chemists often choose easy research topics that can be written up inside a year. There are still some chemistry projects that last five years, but they are rare. Some topics are finished and written up inside six months. It is not unusual for a professor to publish ten papers in a year. And the outcome of a single project is usually split into several parts to produce more papers, which individually offer less information to readers.

I realize that such pressures are common in academia everywhere, but they are most serious in China now, because of the rigid reliance on impact factors, and they are worst in chemistry, because, in chemistry, it is much easier to produce a scientific paper from limited work. Often, papers do not describe truly new chemical reactions, just variations on an established theme. Someone publishes a new reaction, and some chemists just introduce a 'novel' substituent group and run the experiment under 'novel' conditions. The predecessors open a new door and some chemists follow them through wearing different 'novel' clothing.

One way to improve the situation, without slipping back to the days when research funding was allocated to favourites and friends, would be to judge scientists in China not just on journal impact factors but on the citations their papers receive after, say, two years, to see whether their work stands the test of time.

It is right to push Chinese scientists to publish in international journals, but currently there is too much emphasis on presentation and fancy concepts, and too little on solid, high-quality work. Texts that teach scientists how to write world-class papers — with a good structure including among other things, an abstract and an introduction — are popular in China at the moment. But too often these perfectly packaged and neatly written papers discuss empty concepts.

Additional pressure on chemistry in China comes from an increasing emphasis on making research projects 'useful'. China has a big population and lacks resources, so it is right to emphasize the importance of applied chemistry research, which needs an element of process engineering. But science is not directly equal to productivity. Pure chemistry, which explores the unknown and provides new tools for scientists, must not be overlooked.

We are now more than halfway through 2011, the Year of Chemistry. If China is to truly celebrate the discipline and produce good chemistry rather than lots of papers, then change needs to come quickly. ■

IF A HIGH **IMPACT FACTOR** IS THE ONLY GOAL OF **CHEMISTRY** RESEARCH, THEN CHEMISTRY IS NO LONGER **SCIENCE**.

➲ NATURE.COM
Discuss this article online at:
go.nature.com/o5rmeh

**Nai-Xing Wang** *is a professor in the Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, Beijing.*
*e-mail: nxwang@mail.ipc.ac.cn*



© 2011 Macmillan Publishers Limited. All rights reserved

# SEVEN DAYS *The news in brief*

POLICY

## Safer fracking

The natural-gas industry should disclose more information to the public about its shale-gas operations, according to an 11 August draft report from a seven-member advisory panel set up by the US Department of Energy. Disclosure would include greater transparency on the substances used in hydraulic fracturing, or 'fracking', in which high-pressure fluids are pumped into shale to fracture the rock and force out natural gas. The technique has been accused of releasing methane into water and of polluting groundwater with toxic chemicals. See go.nature.com/rbekiq for more.

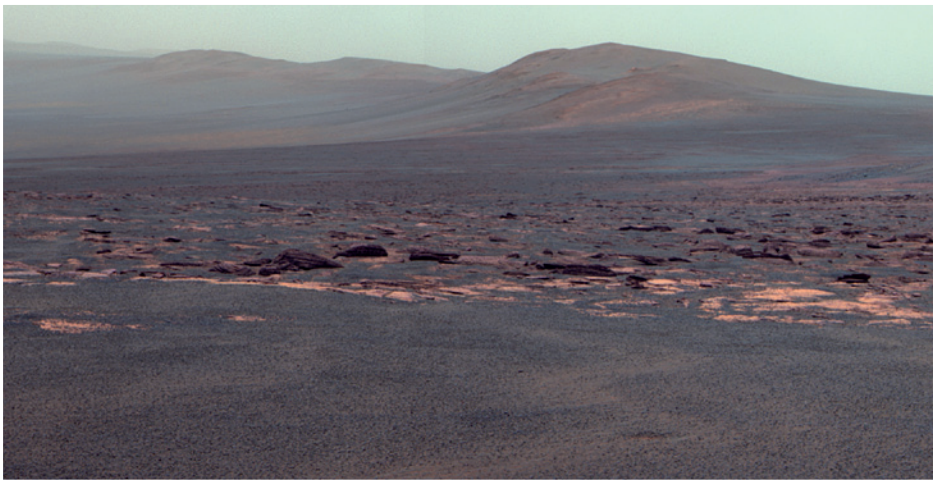
## Chemists campaign

Some 100 senior chemists around the world — including six Nobel laureates — have sprung to the defence of synthetic organic chemistry in the United Kingdom, which is to be squeezed of funding by the nation's Engineering and Physical Sciences Research Council. In letters delivered to British Prime Minister David Cameron on 15 August, chemists say that the cuts would injure the economy and biomedical research. They also cast doubt on the evidence that the council used to justify its decision. See go.nature.com/lsxwhb for more.

NUMBER CRUNCH

# 437

**The record for the most days in space, held by Russian cosmonaut Valeri Polyakov. It was 'broken' on 14 August by the six crew members of the simulation mission Mars500 at the Institute of Biomedical Problems in Moscow. They are mimicking the isolation of an expedition to Mars (without the weightlessness).**



NASA/JPL-CALTECH/CORNELL/ASU

## Mars rover surveys new home

Trekking slowly but steadily across Mars, NASA's Opportunity rover last week completed a three-year journey to reach the rim of a massive crater named Endeavour. At 22 kilometres in diameter, the new site dwarfs the rover's previous home, the 750-metre-wide Victoria crater, which is 21.5 kilometres away. Observations from orbit have shown that Endeavour contains clay minerals, which form under wet conditions, so the site should provide further clues to Mars's wetter past. Opportunity's location at Cape York ridge on the west of Endeavour (pictured) has been named Spirit Point after its twin rover, Spirit, which, after its own 7.7-kilometre expedition, now lies dead on the other side of Mars.

## Tsunami forecast

The Japan Meteorological Agency (JMA) will no longer forecast the height of tsunami waves generated by earthquakes of magnitude 8.0 or above. Instead, it will warn of "the possibility of a huge tsunami" — to avoid any false reassurance. The proposed change comes in response to the early forecast for the Tohoku earthquake on 11 March, which underestimated the size of tsunami waves because the quake continued for two minutes after the JMA took seismic readings. The agency will start discussions of a more thorough overhaul to its tsunami early-warning system in September. See go.nature.com/gsmsck for more.

## Mouse lab reprieve

A laboratory has been spared litigation after it was sued for distributing mouse models of Alzheimer's disease to researchers. In February 2010, the Alzheimer's Institute of America (AIA) in Kansas City, Kansas, had sued the Jackson Laboratory of Bar Harbor, Maine, among other institutes and companies (see *Nature* **472,** 20; 2011). The lab was distributing mouse strains that carried a specific DNA sequence for which the AIA owns the patent, it said. But on 9 August, the AIA dropped its lawsuit after the US National Institutes of Health in Bethesda, Maryland, stepped in to grant the Jackson Lab authorization to distribute the models. Litigation against other defendants continues. See go.nature.com/ceshcc for more.

RESEARCH

## Neutrino spotting

A neutrino-detection experiment in China has started to gather data. Located hundreds of metres underground on the Dapeng peninsula near Hong Kong, the Daya Bay Reactor Neutrino Experiment watches electron antineutrinos produced by two nearby nuclear power plants as they change ('oscillate') into other types of antineutrinos. The


© 2011 Macmillan Publishers Limited. All rights reserved

S. SHOSTAK/SPL



experiment, an international collaboration led jointly by China and the United States, should help scientists to understand more about how neutrinos change type. By comparing oscillations of neutrinos with those of antineutrinos, it might also provide clues about the imbalance of matter and antimatter in the Universe.

## Mach-ed up plane

For the second time, researchers at the US Defense Advanced Research Projects Agency (DARPA) in Arlington, Virginia, have run into problems testing a hypersonic aircraft designed to travel at 20 times the speed of sound (Mach 20, or more than 6,800 metres per second) so that it could swiftly strike anywhere on the planet. About 9 minutes after the arrow-shaped Falcon Hypersonic Technology Vehicle 2 separated from its launch rocket on 11 August, researchers lost contact with the plane, which probably crashed into the Pacific Ocean. A similar fate had befallen the first test, on 22 April 2010.

## Alien hunt restarts

The Allen Telescope Array in northern California (**pictured**) will resume scanning the skies for signs of extraterrestrial life, after a fund-raising appeal to wake the array's 42 radio dishes from hibernation reached its first goal. The array, run by the University of California, Berkeley, and the SETI Institute in Mountain View, California, was shut down in April owing to a lack of funding (see *Nature* **475,** 442–444; 2011). More than US$200,000 has now been raised through SETIstars.org, the SETI Institute announced last week. Keeping the array going will require more money; the institute is hoping to raise it by partnering with the US Air Force to track space debris.

## Epidemics study

This year's outbreak of *Escherichia coli* in northern Europe, which infected 3,910 people and resulted in 46 deaths, has prompted the European Union to allocate an additional €12 million (US$17.3 million) for research into epidemics of infectious diseases. The ANTIGONE network will fund 14 research groups across seven countries to examine how pathogens — such as Shiga-toxin-producing *E. coli*, and Ebola virus — spread from animals or the environment to humans. The project had originally lost out for funding to a similar project, PREDEMICS, focusing on viruses; now both will be funded. See go.nature.com/gfrmiz for more.

BUSINESS

## Emissions network

A private US company that is installing stations around the world to track greenhouse-gas emissions will partner with scientists who are setting up a similar effort in Europe. Earth Networks, based in Germantown, Maryland, is already working with the Scripps Institution of Oceanography in La Jolla, California, as it invests US$25 million over five years in an initial network of 100 monitoring stations. Data will be freely available to researchers. On 16 August, the firm said that it would partner with the European Union's Integrated Observation System (ICOS) to ensure that its stations in Europe are complementary to ICOS's efforts (see *Nature* **465,** 18–19; 2010).

## Stem-cell patent

The US Patent and Trademark Office has approved its first patent for technology used to make induced pluripotent stem (iPS) cells, developed by Shinya Yamanaka of Kyoto University, Japan. The patent was issued on 5 August. A university representative, Akemi Nakamura, said that the university would not exploit the patent to restrict research using iPS cells for non-profit purposes, but that a university spin-off company, iPS Academia Japan in Kyoto, might ask for a licensing fee from commercial companies that use iPS cells. Kyoto University already has patent rights for iPS cell technology in Europe and elsewhere. See go.nature.com/rnumfz for more.

## Nuclear costs

Japan's Tokyo Electric Power Company (TEPCO) has recorded heavy losses in the quarter following the meltdown of three reactors at its Fukushima Daiichi power plant. On 9 August, TEPCO announced a ¥571.8-billion (US$7.4-billion) net loss in the first quarter of the 2011–12 financial year. The loss comes on top of a ¥1.247-trillion hit from the previous financial year, which ended on 31 March. At the power plant, meanwhile, workers are preparing to erect a fabric cover over the Unit 1 reactor to limit the spread of radiation.

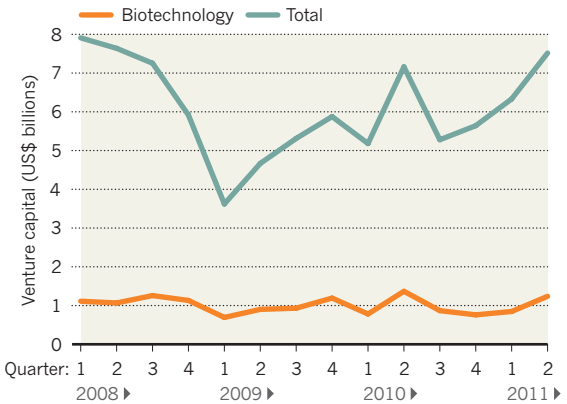
NATURE.COM
For daily news updates see:
www.nature.com/news

## TREND WATCH

Venture capitalists gave US biotechnology firms less money in the second quarter of this year than in the same quarter last year (see chart), according to a report from PricewaterhouseCoopers and the US National Venture Capital Association. The amount invested declined by 9% to US$1.2 billion, and the number of deals dropped by 24% to 116. But investment in medical device companies rose by 9% to $841 million, and investment across all industries rose by 5% to $7.5 billion.

### BIOTECH LAGS IN US VENTURE FUNDING RISE

Although US venture-capital investment is on the rise, the biotechnology sector has yet to return to its 2010 peak.




SOURCE: NVCA/PWC/THOMSON REUTERS



© 2011 Macmillan Publishers Limited. All rights reserved

# NEWS IN FOCUS

**PUBLIC HEALTH** Will HIV prevention pills boost risky behaviour? **p.260**

**POLICY** US agencies promise to protect scientists from political interference **p.262**

**ETHICS** Declared 'flawed' a decade ago, infant-death papers live on **p.263**



**ANIMAL BEHAVIOUR** Surveying chimps for signs of culture **p.270**

G. VAIDYANATHAN



R. BALEIA/LATINCONTENT/GETTY IMAGES

**Soya bean farming in the Amazon is rising along with food prices, contributing to an increase in the rate of deforestation for the first time in years.**

ENVIRONMENT

# Brazil revisits forest code

*Changes to legislation could undermine authorities' power to halt deforestation.*

**BY JEFF TOLLEFSON**

A tough-minded law has boosted Brazil's environmental record in recent years by helping to drive the rate of destruction in the Amazon rainforest to historic lows. But a backlash in the hinterlands is threatening to weaken the country's forest code and push deforestation rates back up again.

Brazil's National Institute for Space Research (INPE) began reporting spikes in deforestation in March, two months before the nation's House of Representatives passed a bill to change the law that governs forests, which has been in force since 1965. The legislation includes clauses supported by agribusiness and small landowners that scale back federal authority and grant immunity from prosecution for all deforestation before 2008. The bill is now pending in the Senate — and is opposed by the administration of President Dilma Rousseff — but some observers say that the debate alone has emboldened landowners to clear more land.

The tensions have been fuelled by high prices for commodities such as soya beans and beef, which have driven up demand for arable land while giving farmers and ranchers the cash they need to bring in chainsaws and bulldozers, says Dan Nepstad, a US ecologist who works with the Amazon Environmental Research Institute in Brasília. Meanwhile, he says, authorities are stuck trying to enforce a law that is strict, complex and difficult to comply with. "What Brazil has created is in many ways sort of the perfect recipe for a rural insurrection," Nepstad says.

But the Rousseff administration is not backing down. Deforestation causes some 15% of global greenhouse-gas emissions and 75% of Brazil's. Rousseff's predecessor, Luis Inácio Lula da Silva, promised during the 2009 climate talks in Copenhagen that Brazil would cut deforestation by 80% by 2020. Although Rousseff is not known as an environmentalist, she has maintained Lula's pledge and sent hundreds of law-enforcement officers into the Amazon as deforestation spiked.

**NATURE.COM**
For more on agriculture in the Amazon, see:
go.nature.com/dpkkvr

The world will get its first glimpse of the results later this month, when the INPE releases its preliminary analysis of the annual deforestation season that ended in July. Some analyses already indicate that the stepped-up enforcement has helped. According to an analysis of satellite data by the non-profit research institute Imazon in Belém, deforestation was actually 42% lower this June than last.

"Deforestation is going to increase in 2011, but it won't be a big jump," says Carlos Souza, a remote-sensing scientist who heads Imazon's satellite analysis. Imazon says that cumulative deforestation between August 2010 and ▶



© 2011 Macmillan Publishers Limited. All rights reserved

June 2011 rose by 15% compared with the same period the season before. If that number holds up in the final analysis using higher resolution satellite data, it would still be the second lowest loss on record (see 'Clearing the Amazon'). "We could say that the situation is still under control," says Souza.

Many scientists fear that the new law will cause further backsliding. For now, all landowners in the Amazon must maintain forest on 80% of their land, and those who had cleared illegally must reforest to that level. The House bill, however, would create exemptions for small landowners, and allow state governments to adjust the 80% rule.

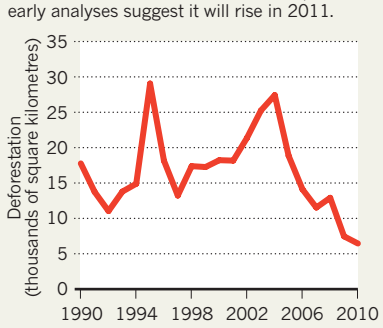
All told, the changes could legalize the clearing of more than 220,000 square kilometres — an area nearly the size of the United Kingdom — according to an analysis by researchers at the University of São Paulo. Thomas Lovejoy, a tropical ecologist at the Heinz Center in Washington DC, says that the bill is a "recipe for Amazon dieback", a feedback loop that decreases rainfall and could convert vast swathes of rainforest into savannah.

But the bill could still be reformed in the Senate. Nepstad is pushing for an affordable way for older violators to comply, and for a credible punishment for new offenders. He says that the government needs to start investing in schemes to protect forest carbon and rewarding frustrated landowners who have obeyed the law while their neighbours cleared more land. If the law can't be enforced, he adds, the alluring profits of agriculture could lead to a "new wave" of deforestation.

Although rural interests have dominated the debate so far, most people share the concerns of the scientific community, says Luiz Martinelli, an ecologist at the University of São Paulo in Piracicaba who is analysing the legislation for senators. A national poll conducted in June by environmental groups found that 85% of Brazilians rate forest protection above agricultural production. "Brazilian society is kind of sick of this deforestation drama," Martinelli says. ■

**CLEARING THE AMAZON**

Law enforcement has helped reduce deforestation in the Brazilian Amazon, but early analyses suggest it will rise in 2011.




AMAZON ENVIRONMENTAL RESEARCH INSTITUTE/INPE



PARTNERS PREP STUDY

**Public-health workers in Thika, Kenya, perform a skit to advertise a prophylactic HIV drug trial.**

DRUG DEVELOPMENT

# HIV drug-prevention strategy carries risks

*Human behaviour complicates prophylactic measures.*

**BY ERIKA CHECK HAYDEN**

When the US Food and Drug Administration approved Viagra in 1998, officials never considered one possible side effect of the drug: higher rates of sexually transmitted diseases among men who, thanks to Viagra, would become more sexually active. A powerful tool in the fight against HIV is raising similar questions about the possibility of unintended public-health consequences if drugs are approved for use in healthy people and cause them to alter their behaviour.

Several studies in the past year have reported that the very drugs used to treat people with HIV can also stop healthy people from becoming infected (see table). But people taking the drugs may adopt riskier behaviours because they feel protected — a phenomenon known as 'risk disinhibition' — undermining the benefit of the drugs and potentially infecting others. Moreover, those who become infected while taking the preventive regimen might develop drug-resistant viruses that they could then transmit to others. "You have this wonderful scientific breakthrough," says Kevin Frost, chief executive of the Foundation for AIDS Research in New York City. "But what are the practical implications?"

Researchers will mull over these issues on 19 August at a meeting convened by the Forum for Collaborative HIV Research in Washington DC. The questions have become more urgent since January, when the drug firm Gilead of Foster City, California, announced that it plans this year to ask the Food and Drug Administration (FDA) to approve its HIV drug Truvada for use in healthy people — in what is known as pre-exposure prophylaxis, or PrEP. Truvada, which contains the antiretroviral drugs tenofovir and emtricitabine, has been used in many of the PrEP trials. In the three clinical trials that have reported benefits for PrEP so far, once-a-day pills have cut a person's risk of acquiring HIV by between 44% and 73%, a variation that is due primarily to differences in how strictly patients stuck to the daily regimen. Although the need for PrEP is greatest in poor countries, approval in the United States could greatly expand the market for Truvada, which generated US$2.65 billion in sales last year.

***"You have this wonderful scientific breakthrough. But what are the practical implications?"***

But asking the FDA to evaluate questions about risk disinhibition and drug resistance might push the agency into uncharted territory. "When you're talking about a population issue, is that something that the FDA should be looking at at all?" asks Jur Strobos, deputy director of the Forum for Collaborative HIV Research.

The clinical trials don't offer clear guidance. Some of the successful trials found that people on PrEP actually used condoms more



© 2011 Macmillan Publishers Limited. All rights reserved

June 2011 rose by 15% compared with the same period the season before. If that number holds up in the final analysis using higher resolution satellite data, it would still be the second lowest loss on record (see 'Clearing the Amazon'). "We could say that the situation is still under control," says Souza.

Many scientists fear that the new law will cause further backsliding. For now, all landowners in the Amazon must maintain forest on 80% of their land, and those who had cleared illegally must reforest to that level. The House bill, however, would create exemptions for small landowners, and allow state governments to adjust the 80% rule.

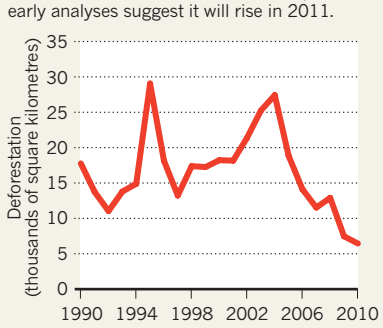
All told, the changes could legalize the clearing of more than 220,000 square kilometres — an area nearly the size of the United Kingdom — according to an analysis by researchers at the University of São Paulo. Thomas Lovejoy, a tropical ecologist at the Heinz Center in Washington DC, says that the bill is a "recipe for Amazon dieback", a feedback loop that decreases rainfall and could convert vast swathes of rainforest into savannah.

But the bill could still be reformed in the Senate. Nepstad is pushing for an affordable way for older violators to comply, and for a credible punishment for new offenders. He says that the government needs to start investing in schemes to protect forest carbon and rewarding frustrated landowners who have obeyed the law while their neighbours cleared more land. If the law can't be enforced, he adds, the alluring profits of agriculture could lead to a "new wave" of deforestation.

Although rural interests have dominated the debate so far, most people share the concerns of the scientific community, says Luiz Martinelli, an ecologist at the University of São Paulo in Piracicaba who is analysing the legislation for senators. A national poll conducted in June by environmental groups found that 85% of Brazilians rate forest protection above agricultural production. "Brazilian society is kind of sick of this deforestation drama," Martinelli says. ■

**CLEARING THE AMAZON**

Law enforcement has helped reduce deforestation in the Brazilian Amazon, but early analyses suggest it will rise in 2011.




AMAZON ENVIRONMENTAL RESEARCH INSTITUTE/INPE



PARTNERS PREP STUDY

**Public-health workers in Thika, Kenya, perform a skit to advertise a prophylactic HIV drug trial.**

DRUG DEVELOPMENT

# HIV drug-prevention strategy carries risks

*Human behaviour complicates prophylactic measures.*

**BY ERIKA CHECK HAYDEN**

When the US Food and Drug Administration approved Viagra in 1998, officials never considered one possible side effect of the drug: higher rates of sexually transmitted diseases among men who, thanks to Viagra, would become more sexually active. A powerful tool in the fight against HIV is raising similar questions about the possibility of unintended public-health consequences if drugs are approved for use in healthy people and cause them to alter their behaviour.

Several studies in the past year have reported that the very drugs used to treat people with HIV can also stop healthy people from becoming infected (see table). But people taking the drugs may adopt riskier behaviours because they feel protected — a phenomenon known as 'risk disinhibition' — undermining the benefit of the drugs and potentially infecting others. Moreover, those who become infected while taking the preventive regimen might develop drug-resistant viruses that they could then transmit to others. "You have this wonderful scientific breakthrough," says Kevin Frost, chief executive of the Foundation for AIDS Research in New York City. "But what are the practical implications?"

Researchers will mull over these issues on 19 August at a meeting convened by the Forum for Collaborative HIV Research in Washington DC. The questions have become more urgent since January, when the drug firm Gilead of Foster City, California, announced that it plans this year to ask the Food and Drug Administration (FDA) to approve its HIV drug Truvada for use in healthy people — in what is known as pre-exposure prophylaxis, or PrEP. Truvada, which contains the antiretroviral drugs tenofovir and emtricitabine, has been used in many of the PrEP trials. In the three clinical trials that have reported benefits for PrEP so far, once-a-day pills have cut a person's risk of acquiring HIV by between 44% and 73%, a variation that is due primarily to differences in how strictly patients stuck to the daily regimen. Although the need for PrEP is greatest in poor countries, approval in the United States could greatly expand the market for Truvada, which generated US$2.65 billion in sales last year.

***"You have this wonderful scientific breakthrough. But what are the practical implications?"***

But asking the FDA to evaluate questions about risk disinhibition and drug resistance might push the agency into uncharted territory. "When you're talking about a population issue, is that something that the FDA should be looking at at all?" asks Jur Strobos, deputy director of the Forum for Collaborative HIV Research.

The clinical trials don't offer clear guidance. Some of the successful trials found that people on PrEP actually used condoms more



© 2011 Macmillan Publishers Limited. All rights reserved

SOURCE: AIDS VACCINE ADVOCACY COALITION

**RISKY BUSINESS**

As clinical trials begin to offer hope that drugs can prevent HIV infections, researchers need to work out how to distribute the drugs without undermining their benefits.

| Trial name | Location | Population | Status/results expected |
|---|---|---|---|
| iPrEx | Brazil, Ecuador, Peru, South Africa, Thailand, United States | 2,499 men who have sex with men (MSM) and transgender women | Ended: benefit reported in November 2010 |
| FEM-PrEP | Kenya, South Africa, Tanzania, Zimbabwe | 3,900 heterosexual women | Ended April 2011: no benefit reported |
| Partners PrEP | Kenya, Uganda | 4,700 heterosexual couples with one infected partner | Ended July 2011: benefit reported |
| TDF2 | Botswana | 1,200 heterosexual men and women | Ended July 2011: benefit reported |
| CDC 4370 | Thailand | 2,400 injecting drug users | First quarter 2012 |
| VOICE | South Africa, Uganda, Zimbabwe | 5,000 heterosexual women | First quarter 2013 |
| ADAPT | South Africa, Thailand | 180 MSM and 180 heterosexual women | In development |
| HPTN 069 | United States | 400 MSM | In development |
| IAVI E002 | Uganda | 72 heterosexual men and women | Ongoing |

frequently while receiving PrEP treatment, countering the risk-disinhibition argument. And only a few instances of drug resistance occurred, and these did not compromise patients' treatments.

But the controlled setting of a clinical trial, in which participants received intensive prevention counselling and were tested monthly for HIV, is very different from the real world. "We think of PrEP as a pill, but we all recognize that PrEP is about a much broader programme," says Mitchell Warren, director of the AIDS Vaccine Advocacy Coalition in New York City. "How you would deliver the kind of testing and monitoring that would go into that programme is a very important question."

The US Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, has recommended testing patients for HIV before they begin taking the drugs and again at two- to three-month intervals. The FDA could also require drug companies to set up a registry of patients taking the drugs and ask that patients provide proof of a negative HIV test before getting their medications refilled.

Deciding who to treat with PrEP could also be a challenge. The CDC reported on 3 August that rates of new HIV infection in the United States are stable overall, but are rising in young men who have sex with men. Yet if these men aren't using the prevention measures already available, there's little reason to think doctors will have an easier time convincing them to take a daily pill.

The question of who should get PrEP is more difficult in many developing nations, which cannot even afford to treat everyone currently infected with HIV. PrEP would cost hundreds of dollars per patient per year in developing countries, and many thousands of dollars in rich nations.

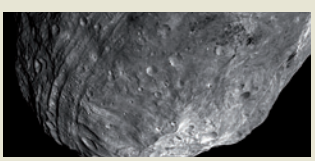
Even with regimens costing less than a dollar a day, developing nations will be forced to choose between providing more treatment for those who already need it and potentially preventing new infections. Myron Cohen, a doctor and researcher at the University of North Carolina at Chapel Hill, points out that half of young girls in some parts of sub-Saharan Africa become infected with HIV by their mid-twenties. "That's unacceptable, so I see that as one potential population for PrEP," says Cohen.

Although the cost-effectiveness of PrEP increases in higher-risk populations, it will be politically dicey for financially strapped countries to justify distributing drugs to those in these groups. "Even if you thought the best use of the pills would be for sex workers, it would be very difficult to take a limited supply of pills and give them to high-risk populations at the expense of people who are dying of infection," says Cohen. ■

NASA/JPL-CALTECH/UCLA/MPS/DLR/IDA

**MORE ONLINE**

**TOP STORY**



Astronomers investigate strange hole on Vesta go.nature.com/jdomt8

**OTHER NEWS**

- California readings shed light on spread of Fukushima's radioactivity go.nature.com/ucld8a
- Mexican marine protected area is 'the best in the world' go.nature.com/r1a716
- 'Electronic skin' could replace bulky electrodes go.nature.com/gftxum


© 2011 Macmillan Publishers Limited. All rights reserved



REUTERS/US COAST GUARD

US agencies initially challenged scientists' estimates of oil plumes from the 2010 Deepwater Horizon spill.

POLICY

# Agencies unveil plans to safeguard science

*White House is reviewing 19 'scientific integrity' policies submitted by US agencies.*

BY EUGENIE SAMUEL REICH

Microbiologist David Lewis knew he might upset the biosolids industry with his research, which suggested that the spreading of sewage sludge on land could make people sick. But he didn't expect his employer, the US Environmental Protection Agency (EPA), not to back him up.

Following complaints from a biosolids company, an EPA official used agency letterhead to spread allegations of research misconduct against Lewis, a situation that Lewis says led to him being forced out in 2003. A judge in 2007 dismissed Lewis's claims that he was fired illegally, but the EPA cleared Lewis of misconduct and reprimanded the official. With his science career nevertheless over, Lewis now works at the National Whistleblowers Center in Washington DC.

Lewis's fate is the kind that a scientific-integrity policy released by the EPA on 5 August seems intended to prevent. The agency is one of 19 that have submitted policies to the White House Office of Science and Technology Policy (OSTP), although the policies are not all complete. At least six have been released publicly. Developed after a 2009 memo in which US President Barack Obama sought to distinguish his administration's use of science from its alleged political abuse under his predecessor George W. Bush, the policies are meant to put sound science at the centre of government policy-making.

Many scientists endorse the goal, but Lewis is among those who are cynical about what they have seen so far. "There's nothing in EPA's draft policy that even hints at dealing with things done to scientists like me to shut us up," he says. He adds that the agency needs new people with good records in scientific integrity, not new policies.

An EPA spokeswoman declined to comment on Lewis's views, but notes that the policy is a draft, open for public comment and subject to revision. One agency, the Department of the Interior, has made public a final policy, and NASA says it will revise existing policies rather than offer a new one. The Department of Energy and the Department of Health and Human Services (which includes the National Institutes of Health) have yet to publish drafts.

Some scientists say that the new policies will be beneficial. Michael McPhaden, an oceanographer at the National Oceanic and Atmospheric Administration (NOAA) and president of the American Geophysical Union, likes the fact that the NOAA policy encourages media interviews. "The agency is now encouraging you to speak publicly about your research instead of putting roadblocks in your way," he says.

But critics worry that some policies lack specificity. The watchdog group Public Employees for Environmental Responsibility, based in Washington DC, praises the policies of NOAA and the Department of the Interior, which explicitly ban politically motivated alteration of technical documents, but it notes that such a clause is missing from the EPA draft (see table). Neal Lane, former director of both the OSTP and the National Science Foundation, is also concerned about an opaque reference to "clearance procedures" in the EPA draft. Although all of the policies released so far allow employees to speak publicly, the EPA's phrase suggests that there can still be scope for scientists to be muzzled, says Lane, a physicist at Rice University in Houston, Texas.

Roger Pielke Jr, a science-policy expert at the University of Colorado, Boulder, says that differences in the policies are inevitable because the agencies have different missions and scientific integrity is a vague concept. OSTP spokesman Rick Weiss says the office intentionally allowed agencies to respond to the effort in different ways. The OSTP expects to work with agencies to finalize the policies by this autumn.

Lane says that, whatever their flaws, the policies are a step forward. "Having these policies on the books will make managers think twice before interfering with a scientist," he says. ■ SEE EDITORIAL P.251

**KEEPING POLITICS OUT OF SCIENCE**

US agencies have drafted policies to protect science from political interference, but five policies made public so far differ in the details.

| Agency | Scientists can voice opinions publicly/ speak to media | Bans politically motivated alteration of technical documents | Explicitly encourages publication of scientific findings |
|---|---|---|---|
| Department of the Interior | Yes, under unspecified "procedures" | Yes | No |
| National Oceanic and Atmospheric Administration | Yes, if clear they do not speak for the agency | Yes | Yes |
| Environmental Protection Agency | Yes, if following "clearance procedures" | No | Yes, subject to management approval |
| National Science Foundation | Yes, if clear they do not speak for the agency | No. Science is conducted externally | Yes |
| NASA | Yes, but must notify manager. Encouraged to have press officer present | No | Yes |



© 2011 Macmillan Publishers Limited. All rights reserved

ETHICS

# 'Flawed' infant death papers not retracted

*Authors, institutions and editors did not act on findings of inquiry into organ harvesting.*

BY JO MARCHANT

"Fundamentally flawed": that's how a January 2001 report described ten research papers based on body parts taken from hundreds of dead children without their parents' consent. Yet more than a decade on, *Nature* has learned, only one of those papers has been removed from the scientific record.

The lack of action, even in what seems to be a clear-cut case, highlights the reluctance of institutions and journals to retract papers when the authors stand by the results.

A series of inquiries by the government, the police and the General Medical Council (GMC) revealed that until the late 1990s, pathologists in the United Kingdom had routinely been taking organs and other samples from children's bodies without asking parents for permission. The findings caused shock and disbelief across the country, and led to an overhaul of human-tissue legislation.

The scandal centred on Alder Hey Children's Hospital in Liverpool. A government-commissioned report chaired by the lawyer Michael Redfern concluded that Dutch pathologist Dick van Velzen routinely removed every organ from the bodies that he examined at the hospital, storing the tissues for research. Yet van Velzen often did not carry out proper post-mortems on these children, filing incomplete autopsy reports or making them up completely.

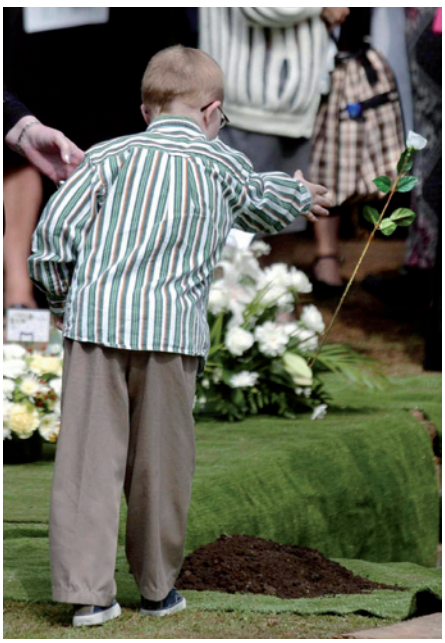
Van Velzen published several papers on these tissue samples with collaborators at the University of Liverpool, where he held a post. Much of this work aimed to investigate the causes of unexplained cot death, or sudden infant death syndrome (SIDS), by comparing infants who had died from SIDS with those who had died from other causes.

The Redfern report concluded that because van Velzen did not carry out proper post-mortems, it was impossible to be sure how any of the infants — in either the SIDS group or the control group — had actually died. In the light of this and other problems, the report highlighted ten peer-reviewed papers[1–10] that it concluded were "fundamentally flawed". The GMC subsequently banned van Velzen from practising medicine in the United Kingdom, because of his serious professional misconduct in the matter. No action was taken against van Velzen's co-authors. *Nature* has been unable to contact van Velzen.

In 2009, Christian Holscher, a neuroscientist at the University of Ulster, decided to find out what had happened to the papers. Vyvyan Howard, a colleague of van Velzen's at Alder Hey and co-author on all ten papers, was by then working in Holscher's department.

Holscher was surprised to find that none of the papers had been retracted. "It's not okay that these papers are still in the public domain," he says. "If you don't know the background, you assume this is a scientific study with proper results."

Holscher wrote to the editors of several journals involved to ask why the papers had not been retracted. Victor Chernick of *Pediatric Pulmonology* retracted the paper in his journal[3] almost immediately — so far, it is the only paper on the list to have been withdrawn. Elia Maalouf of *Early Human Development* replied to say that he would retract that journal's paper[5], but this has not happened. Miguel Reyes-Múgica of *Pediatric and Developmental Pathology* said he saw no need to retract the paper in his journal[4].



C. LANE/REUTERS/CORBIS

**Paying respect at the burial of infant remains that came to light as a result of the Alder Hey inquiry.**

Reyes-Múgica told *Nature* that it is not up to him to judge whether the paper is invalid, despite the findings of the Redfern report. "The editor's job does not have to do with judicial matters," he says.

Five of the ten papers are currently available on PubMed[2,4–7], and three are included on the University of Ulster's online publication database[2,4,5]. Most of the papers have not been highly cited, but one[7] — which shows that growth-retarded fetuses have poor kidney development — receives a steady stream of PubMed citations, and has been cited six times since 2010.

## CLINICAL CONTRIBUTION

Howard says that he and his co-authors considered whether to retract their papers after the Redfern report was published, but ultimately decided that all of their research remains valid. "We think we've made a contribution there," he says. "The outcome of that research has affected clinical practice," Howard says, in that growth-retarded fetuses are now delivered as soon as possible to improve their chances of survival.

He and the other co-authors argue that Redfern and his colleagues misunderstood their work. Howard and Paul Sibbons, now director of operations at Northwick Park Institute for Medical Research in Harrow, UK, and co-author on seven of the criticized papers, say that several papers do not rely on a comparison between SIDS and non-SIDS cases, and so the lack of proper diagnosis would not affect the conclusions. Sibbons adds that even in the studies that did rely on a comparison, his analysis of the tissue samples provided diagnoses that were adequate for the research. He also says that he was "astounded" that the Redfern report contained "so many inaccuracies", and says that his research should not have been classified as "flawed".

"My findings of fact have remained unchallenged for 11 years," Redfern says. He declined to comment on the authors' decision not to retract the criticized papers. "That is their choice," he told *Nature*. "It is for you, I think, to draw conclusions."

Holscher, who describes the Redfern report as "detailed and meticulous", argues that the University of Liverpool should have requested that the papers be retracted, rather than leaving it up to the authors. "The editors of these journals often had no idea that these papers were found to be invalid," he says. "It's the ▶



© 2011 Macmillan Publishers Limited. All rights reserved

university's job to tell them. It's not in the interests of the authors."

In a statement, a spokeswoman for the University of Liverpool said that the university has strict guidelines and standards for research integrity. "Following publication of the Redfern report, the university prevented further publication of work based on samples that were the subject of the inquiry." She added that the university's Research Governance Committee is now considering the papers that were not retracted. "We will request the papers are withdrawn if we feel this is appropriate."

The three papers currently included on the University of Ulster's online database are "legitimate published output from a member of staff", according to Norman Black, pro-vice-chancellor for research and innovation at the university. "The university is not aware of any request or recommendation by the Alder Hey inquiry to retract the papers," he says.

## TREMENDOUS VARIATION

Ivan Oransky, a US science journalist who runs the blog Retraction Watch, questions the wisdom of leaving retractions either to authors or their institutions, noting that both are likely to be reluctant. He cites the case of Harvard psychologist Marc Hauser, who resigned from his position last month after an internal investigation, which has not been published. According to Harvard, Hauser was guilty of eight counts of scientific misconduct relating to three papers (see *Nature* **466,** 908–909; 2010). "There was only one retraction," says Oransky. "The investigation took Harvard three years, and we still don't know what it says."

Instead, he argues that journal editors should play a more active role. "Editors should stand by what they publish," he says. He complains of "tremendous variation" in the attitudes of journal editors, but cites as a positive example Steven Shafer, editor-in-chief of *Anesthesia & Analgesia*, who led a group of editors of 11 affected journals after the German anaesthesiologist Joachim Boldt was found to have published research that had been conducted without proper ethical approval. Working together, the editors have since retracted around 90 of Boldt's papers.

***"The editors of these journals often had no idea that these papers were found to be invalid."***

Sibbons, who is himself editor of the journal *Comparative Clinical Pathology*, says that for an editor to judge properly whether papers should be retracted in cases such as these would require a retrospective review process, something that is unlikely to be worth the cost. He argues that follow-up experiments already provide a mechanism to check the validity of studies. "Replication of experiments is usually what highlights anomalies, not retraction," he says.

But Oransky points out that according to guidelines published by the Committee on Publication Ethics, a group of journal editors, it is within the scope of the editor to retract a paper if an investigation finds it to be flawed.

When the authors do not agree to the retraction, he suggests that an editor could ask them for a rebuttal and print that as part of the retraction notice. In the absence of a full retraction, he says that a warning label alerting others to any concerns would be "a great thing". The aim, he says, is to increase transparency and "treat science as a process rather than a product". ■


1. Ansari, T. *et al. Eur. J. Morphol.* **33,** 294–298 (1995).
2. Beech, D. J., Howard, C. V., Reed, M. G., Sibbons, P. D. & van Velzen, D. *J. Microsc.* **197,** 36–45 (2000).
3. Beech, D. J., Sibbons, P. D, Howard, C. V. & van Velzen, D. *Pediatr. Pulmonol.* **31,** 339–343 (2001).
4. Beech, D. J., Sibbons, P. D., Howard, C. V. & van Velzen, D. *Pediatr. Dev. Pathol.* **3,** 450–454 (2000).
5. Beech, D. J., Sibbons, P. D., Howard, C. V. & van Velzen, D. *Early Hum. Dev.* **59,** 193–200 (2000).
6. Hinchliffe, S. A. *et al. Pediatr. Pathol.* **13,** 333–343 (1993).
7. Hinchliffe, S. A., Lynch, M. R., Sargent, P. H., Howard, C. V. & van Velzen, D. *Br. J. Obstet. Gynaecol.* **99,** 296–301 (1992).
8. Howard, C. V. *et al. Zool. Studies* **34,** 109–110 (1995).
9. Ricketts, S. A. *et al. J. Cell. Pathol.* **3,** 17–26 (1998).
10. Sibbons, P. D., Ricketts, S. A. & van Velzen, D. *J. Cell. Pathol.* **1,** 153–160 (1996).


© 2011 Macmillan Publishers Limited. All rights reserved

D. GONSEAU



Gudgeon downstream of a wastewater-processing plant had swollen abdomens and other abnormalities.

ENVIRONMENT

# Drug waste harms fish

*Discharges from pharmaceutical factories contaminate rivers on three continents.*

BY NATASHA GILBERT

Consumers who flush unwanted contraceptives down the drain have long been blamed for giving fish more than their fair share of sex organs. Drugs excreted by patients can also taint rivers, even after passing through wastewater-processing facilities.

But evidence is accumulating that the effluent coming from pharmaceutical factories could also be carrying drugs into rivers. Many ecotoxicologists had assumed that water-quality standards, along with companies' desire to avoid wasting valuable pharmaceuticals, would minimize the extent of bioactive compounds released by factories into wastewater, and ultimately into rivers.

A string of studies suggest otherwise. In 2009, for example, researchers[1,2] reported very high levels of pharmaceutical ingredients in treated effluent coming from a plant that processes wastewater from factories near Hyderabad, India. The following year, a similar discovery was made at two wastewater-treatment plants in New York, both of which received discharges from drug-production plants[3].

Now, researchers have provided the first evidence of similar problems in Europe[4], and have linked it to sex disruption in wild fish populations found in the Dore River in France. "People thought this could not happen in a country that has high environmental standards and good manufacturing practices," says Patrick Phillips, head of the National Water-Quality Assessment Program at the US Geological Survey in Troy, New York, and lead author of the US study. "The evidence from the United States and now from France shows that this is not the case."

The discovery has prompted calls for more effective oversight of the industry. The United States, the European Union, the United Kingdom and France do not have regulations limiting the concentrations of pharmaceuticals released into the aquatic environment in either municipal wastewater or in effluent from manufacturing facilities. "People think drug release is regulated, but its not," says Joakim Larsson, a pharmacologist at the University of Gothenburg in Sweden, and an author of one of the Indian studies[2].

The French study investigated the health of wild gudgeon (*Gobio gobio*) populations in a river near a facility close to Vertolaye, owned by pharmaceutical multinational Sanofi, which produces steroid compounds. It was commissioned by the French environment ministry after anglers spotted abnormal fish in the area.

Downstream from the factory, the researchers found that on average 60%, and in one case 80%, of the fish had both male and female sexual characteristics. Upstream of the effluent discharge, such intersex fish made up just 5% of the populations (see 'Fishy find').

Male fish living downstream from the factory also had significantly higher blood levels of vitellogenin, a protein normally found in eggs, than those living upstream.

"This is a real problem," says Wilfried Sanchez, an ecotoxicologist at the French National Institute for Industrial Environment and Risks, and lead author of the study. Sexual abnormalities in gudgeon may not only prevent the fish from breeding, but also signal problems in other species, and a reduction in the fish population could have broader consequences for the river ecosystem.

## MAIN CULPRITS

In results they have yet to publish, Sanchez and his colleagues identified the main pharmaceutical pollutants in the river as being dexamethasone (an anti-inflammatory and immunosuppressant), spironolactone (a diuretic that also blocks the effects of male sex hormones) and canrenone, also a diuretic. All were measured at concentrations of around 10 micrograms per litre, which is "very high" for biologically active substances, says Sanchez. It is unclear how these compounds ended up in the river.

Ian Weatherhead, Sanofi's communications director, points out that because "no effect has been observed in other fish species", it is difficult to be sure how widespread the problem is. He says that the company is cooperating with regulatory agencies, researchers and ecological associations to help identify the cause of the "disturbances", which are "probably multifactorial".

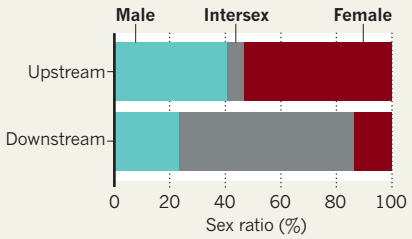
Some scientists and policy-makers say that broader regulations are needed. The European Commission is now considering whether to set limits on the handful of drugs commonly found in waterways, including the anti-inflammatory drug ibuprofen and the contraceptive compound ethinylestradiol. It expects to publish its decision in the autumn as part of proposed reforms to the Water Framework Directive, which governs water pollution. The reform may, for example, demand more extensive cleaning processes in wastewater-treatment plants.

Complicating the regulators' efforts, scientists have not determined safe limits for many pharmaceuticals in the aquatic environment, or how widespread the problem is, says Susan Jobling, an aquatic ecotoxicologist at Brunel University in London. "We don't know the extent to which it is happening," agrees Larsson. "Very few manufacturing sites have been measured." ■

SOURCE: REF. 4

**FISHY FIND**

Fish downstream of a French pharmaceutical factory were much more likely to show characteristics of both sexes (intersex) than those upstream.




1. Gunnarsson, L. *et al. Environ. Toxicol. Chem.* **28,** 2639–2647 (2009).
2. Carlsson, G., Örn, S. & Larsson, D. G. *Environ. Toxicol. Chem.* **28,** 2656–2662 (2009).
3. Phillips, P. J. *et al. Environ. Sci. Technol.* **44,** 4910–4916 (2010).
4. Sanchez, W. *et al. Environ. Int.* **37,** 1342–1348 (2011).



© 2011 Macmillan Publishers Limited. All rights reserved

# The cultured chimpanzees

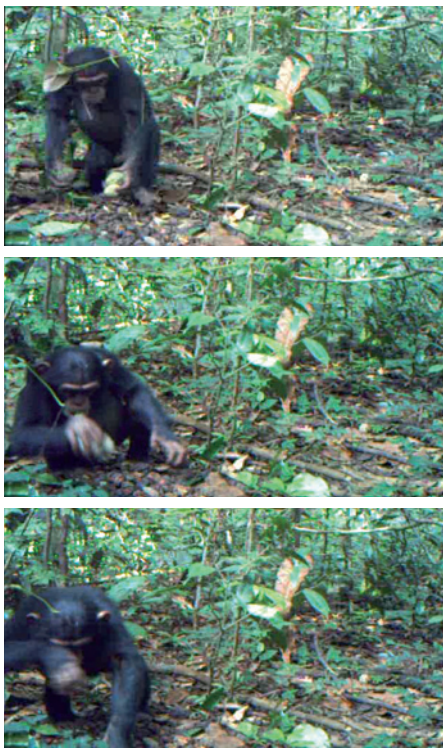
## Do chimpanzees have traditions? As wild populations dwindle, researchers are racing to find out.

BY GAYATHRI VAIDYANATHAN

Thump! Thump! Thump! As the hollow sound echoes through the Liberian rainforest, Vera Leinert and her fellow researchers freeze. Silently, Leinert directs the guide to investigate. Jefferson 'Bola' Skinnah, a ranger with the Liberian Forestry Development Authority, stalks ahead, using the thumping to mask the sound of his movement.

In a sunlit opening in the forest, Skinnah spots a large adult chimpanzee hammering something with a big stone. The chimpanzee puts a broken nut into its mouth then continues pounding. When Skinnah tries to move closer, the chimp disappears into the trees. By the time Leinert and her crew get to the clearing, the animal is long gone.

For the past year, Leinert has been trekking through Sapo National Park, Liberia's first and only protected reserve, to study its chimpanzee population. A student volunteer at the Max Planck Institute for Evolutionary Anthropology (EVA) in Leipzig, Germany, Leinert has never seen her elusive subjects in the flesh but she knows some of them well. There's an energetic young male with a big belly who hammers nuts so vigorously he has to grab a sapling for support. There are the stronger adults who can split a nut with three blows. And there are the mothers who parade through the site with their babies. They've all been caught by video cameras placed strategically throughout Sapo.



In Liberia, camera traps helped researchers to get to know a young chimp, seen here cracking nuts.

MAX PLANCK INST. EVOL. ANTHROPOL.

Chimpanzees in the wild are notoriously difficult to study because they flee from humans — with good reason. Bushmeat hunting and human respiratory diseases have decimated chimpanzee populations[1], while logging and mining have wiped out their habitat. Population numbers have plunged — although no one knows by exactly how much because in most countries with great apes, the animals have never been properly surveyed.

The Pan Africa Great Ape Program, the first Africa-wide great-ape census to be mounted, could change that. In addition to surveying chimpanzee numbers (see 'How many chimpanzees are left?'), project scientists plan to set up automated video and audio recording devices at 40 research sites in 15 countries with chimp populations. Led by Christophe Boesch, director of the primatology department at the EVA, and Hjalmar Kühl, also at the EVA, the programme aims to get a picture of how chimpanzee behaviour — from nut cracking to vocal calls — varies across Africa. Ultimately, the hope is to learn about the origins and extent of what, in humans, would be called culture.


© 2011 Macmillan Publishers Limited. All rights reserved

A. SHAH/NATUREPL.COM



**A female chimpanzee in the Kasakela community in Gombe, Tanzania, shows her offspring how to use a stem as a tool to remove termites from a mound.**

Until recently, scientists regarded culture — defined as socially transmitted behaviours — as exclusive to humans, but there is growing recognition that many animals exhibit some sort of culture. Chimpanzees, which share 98% of their genes with humans, have the most varied set of behaviours documented in the animal world. The difference between humans and animals is growing less distinct, say some researchers. "It is not black and white," says Kühl, who is Leinert's supervisor at the EVA.

In the old scenario, "only humans have culture", says Jason Kamilar, a biogeographer in the department of anthropology at Yale University in New Haven, Connecticut. "Then, culture would be the defining feature of humanity, which evolved some time after the split between the human and chimp lineages," he says. But "if chimps have culture, then presumably the last common ancestor of chimps and humans had culture".

## MAPPING BEHAVIOUR

Some chimps dance slowly at the beginning of rain showers, others don't; some use long sticks to dig up army ants; others use short sticks. In West Africa, some chimp groups hammer nuts with a stone or a piece of wood to open them. But east of the river Nzo-Sassandra, which cuts across Côte d'Ivoire, only one group has been seen cracking nuts.

So far, researchers have observed these variations over years spent studying groups of chimpanzee that have been carefully habituated to the presence of humans. There are just 12 such colonies in Africa (see 'Chimpanzee census'), the most famous of which is in Gombe Stream National Park in Tanzania, where primatologist Jane Goodall worked.

> "THE DIFFERENCE BETWEEN HUMANS AND ANIMALS IS GROWING LESS DISTINCT."

In 1999, evolutionary psychologist Andrew Whiten of the University of St Andrews, UK, and his colleagues compiled a list of behaviours seen in seven of those groups and showed that chimpanzees have unique traditions depending on where they live[2]. They identified at least 39 behaviours from a list of 65 that varied between groups for no obvious reason.

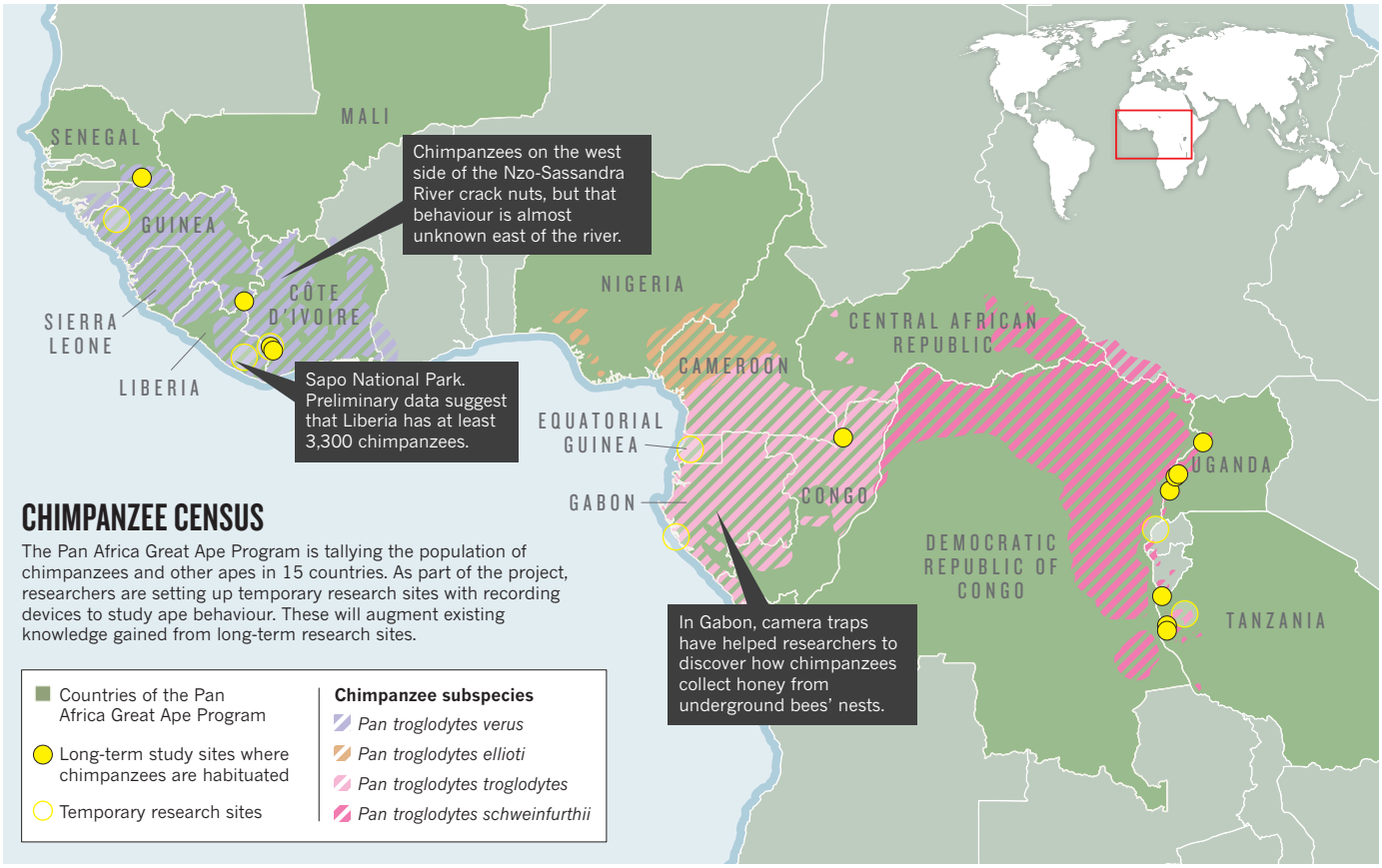
In humans, culture is passed on from one person to another, and in laboratory studies chimpanzees have shown the capacity to pass on learned customs. In one experiment, Whiten and his colleagues taught two chimps a complex series of steps for getting food from a box. Soon after the chimps were reunited with their groups, all the animals were using this method to get their food[3]. But whether such social learning happens in the wild is less clear. Gorillas and bonobos can also learn to use tools in the lab, but rarely use them in their natural habitat[4].

Deciphering culture in the wild is difficult because researchers must ensure that behavioural differences between groups do not have other causes, such as variation in genetics or environmental conditions. "Why is it all chimps don't do everything? One solution is that there are hidden ecological differences between populations," says primatologist Richard Wrangham at Harvard University in Cambridge, Massachusetts. A behaviour could be linked to any number of variables such as amount of rainfall, the types of tree available, or the kinds of predator in the area, he says.

These influences can be subtle, as researchers found while studying how chimps use sticks to harvest army ants. Chimpanzees in



© 2011 Macmillan Publishers Limited. All rights reserved




SOURCE: HJALMAR KÜHL

Guinea sometimes use short sticks and sometimes use sticks up to twice as long. No reason for this was obvious until Tatyana Humle, an anthropologist at the University of Kent, UK, found that some ants are more aggressive, with longer legs and larger mandibles; they run up sticks quicker and bite harder[5]. This might explain why chimps elsewhere in Africa also choose tools of varying lengths to get at ants.

But researchers have not been able to find obvious explanations for other variations related to ant harvesting. Chimpanzees in Cote d'Ivoire sweep the ants off their sticks and into their palms before eating; in Guinea, only about 320 kilometres away, the animals stick the ant-laden sticks directly into their mouths. The same type of ant is present in both places.

Ruling out genetic influences is equally complicated. This year, molecular ecologist Kevin Langergraber at the EVA and his colleagues compared genetic and behavioural data for nine groups of chimpanzee. They found that communities with greater overlap in their mitochondrial DNA showed more similarities in their behaviour[6]. "What we are saying is, you haven't really ruled out the genetic explanation," says Langergraber.

There may be a few hundred thousand chimpanzees in Africa, but researchers have studied just 700–1,000 chimpanzees at the dozen sites with well habituated colonies, says Whiten. The available information from those groups is too little to determine how genes and the environment influence behavioural variations. Kühl compares the situation to using a handful of villages scattered around the world to draw basic conclusions about all the rituals that define human culture.

Whiten and his colleagues are now carrying out more detailed comparisons of the behaviour and ecology of chimps at all the habituated sites. But it has taken 50 years to capture the data they are using, most of which were recorded by painstaking observational studies.

The way forward may be the use of cameras hidden in strategic sites, like those Leinert and her team are setting up in Liberia. Such techniques have already proved their worth. Two years ago in Gabon, Boesch and his team were puzzled by random pits they observed in the ground. They set up camera traps and obtained video recordings of chimps digging to extract honey from underground bees' nests — something that had never been seen before[7]. "Camera traps are proving to be an exciting way to reveal new and often complex behavioural techniques in wild chimpanzee communities," says Whiten.

## CAUGHT IN THE ACT

At the site in Sapo, Leinert pulls on gloves to measure the rock used by the chimp to crack open nuts of the Guinea plum, *Parinari excelsa*. The rock is sizeable, weighing in at 880 grams. She collects nuts for later analysis, as well as hair and dung samples for genetic studies.

Leinert may later put up a video camera at the location to collect more data on the



G. VAIDYANATHAN; V. LEINERT

**Hjalmar Kühl (left, front) and fellow researchers are surveying Liberia's Sapo National Park to tally chimpanzee numbers. Audio (right) and camera traps in the park record chimp behaviour.**


© 2011 Macmillan Publishers Limited. All rights reserved

## HOW MANY CHIMPANZEES ARE LEFT?

*Researchers are trekking across 15 nations to find out.*

Jacob the chimp, now two years old, spends most of his day in a wooden box not much bigger than himself. Born in Sapo National Park in Liberia, he was rescued by a forest ranger, who found Jacob and his dead mother in the arms of a poacher.

Such tales are common in Africa. Bushmeat is a vital source of protein and a dead chimpanzee can fetch US$200 in Nigeria. No one knows exactly how many chimps there are in the wild: in 2003, the International Union for Conservation of Nature made a very rough estimate of 172,700–299,700. But the population is declining rapidly, and many communities are likely to disappear in the next few decades. A study in 2008 found that the population in Côte d'Ivoire had decreased by 90% in 17 years.

In 2010, the dearth of data prompted the Max Planck Institute for Evolutionary Anthropology (EVA) in Leipzig, Germany, to team up with the Wild Chimpanzee Foundation, headquartered at the EVA, and Conservation International, based in Arlington, Virginia, to launch the Pan Africa Great Ape Program. They aim to conduct nationwide surveys in 15 countries to estimate how many chimps are left in Africa. The scientists involved would not disclose the project's budget, but acknowledged that the surveys will be expensive and that they do not yet have all the necessary funding.

As part of the survey, graduate student Jessica Junker of the EVA and her Liberian team of graduate students and rangers from the Forest Development Authority are walking some 400 kilometres to survey 68 squares laid out on a grid across the country. They trek through uncut bush and overgrown farms, across rivers, and into deep muddy valleys to look for chimpanzee nests. Each chimp usually builds a new nest every day, and the researchers can estimate the age of a nest from its state of decomposition. They can then extrapolate to get an idea of the number of animals in an area. Their findings so far suggest that Liberia holds at least 3,300 chimpanzees.

Using similar methods in Sierra Leone, the 2008–10 Tacugama National Chimpanzee Census estimated that more than 5,500 chimpanzees live in that country. This is much higher than a 1981 estimate of 2,500, probably because the earlier survey used less systematic survey methods.

Christophe Boesch of the EVA, who co-heads the Pan Africa Great Ape Program, says that it will guide conservation efforts to where they can do the most good. But getting precise numbers on the great apes in each country is expensive because of the labour involved, and some conservationists would rather see the money spent on enforcing laws against poaching.

"We don't need a nationwide survey to tell us we are losing the battle," says David Greer, who coordinates the African Great Apes Program for the conservation group WWF. "We need to be more assertive, more aggressive with intervention measures, trying to stop the decline." **G.V.**

nut-cracking behaviour. The cameras are mounted in boxes on tree trunks at the height of a chimp's shoulder, and powered by rechargeable batteries. An infrared motion detector activates the camera for one minute when anything moves in its range.

Near the nut-cracking site, a solar-powered audio device is already continuously recording the forest sounds. Chimpanzees emit a range of calls, including short, high-pitched 'pant hoots' that are unique to each individual, and researchers can use them to identify individuals and to tally the size of a community. These calls may be a form of vocal culture, somewhat like human dialects[8].

Over the next five years, the Pan Africa Great Ape Program will establish similar recording stations at locations across Africa. "So potentially we might have, in a few years, behavioural differences from 40 different populations, which is, as you know, four times more than what we have now," says Boesch.

Kühl proposes that these data could help in designing computer models to test how genes, ecology and social transmission influence the distribution and spread of behaviours such as nut cracking. One idea is that when female chimpanzees reach sexual maturity and move to new communities, they pass along their learned behaviours. Another possibility is that each group invents its own behaviours, some of which catch on and become a culture. Individual practices can die out in particular groups but thrive in others. Or, it might be that some chimp groups refuse to take up new ways of doing things from incoming individuals. This could explain why some populations show similar behaviours and others do not.

Before Kühl and his colleagues can conduct the modelling work, they need to devise a faster way to go through the recordings made by the camera and audio traps, which are accumulating at a rate of hundreds of hours each month. Students are currently carrying out the analysis but it can take 10 hours to go through an hour of video, according to Kühl. So engineers at the Fraunhofer Institute for Digital Media Technology, based in Ilmenau, Germany, have developed a computer algorithm to recognize individual chimpanzees from their facial patterns and distinctive features, such as the wrinkles under their eyes. In tests of zoo animals, the software can correctly identify individual chimpanzees 83% of the time, and it processes recordings ten times faster than a person can.

**NATURE.COM**
To view a video of chimp behaviour:
go.nature.com/augwrk

Scan the tag above with the free app from gettag.mobi

Nevertheless, the cameras cannot reveal how an adult chimp patrols its range, or other actions that play out over a wide area. The full portfolio of traditions in the community will remain a mystery. And automated recordings will never capture the subtle ecological information — such as the mandible size and leg length of army ants — that may eventually explain particular behaviours. These require boots on the ground, and long-term behavioural studies are needed to see how chimpanzees pass traditions on to each other as a driver of culture.

But already, the 30 cameras that Leinert has set up in Sapo Park have delivered some tantalizing clues. She is most interested in the lively young male she calls 'Janosch', whom she likes for "his big belly and the way he strikes out to crack the nuts". Besides being entertaining, he sometimes carries his pounding rock away with him, something Leinert hasn't seen with most other chimpanzees in Sapo. The practice may yet catch on with others there. If so, Leinert could be seeing the beginnings of a cultural variation, captured by the cameras she set up in the forest. ■

**Gayathri Vaidyanathan** *is an International Development Research Center fellow at* Nature.

1. Köndgen, S. *et al. Curr. Biol.* **18,** 260–264 (2008).
2. Whiten, A. *et al. Nature* **399,** 682–685 (1999).
3. Whiten, A., Horner, V. & de Waal, F. B. M. *Nature* **437,** 737–740 (2005).
4. McGrew, W. C. *Science* **328,** 579–580 (2010).
5. Schöning, C., Humle, T., Möbius, Y. & McGrew, W. C. *J. Hum. Evol.* **55,** 48–59 (2008).
6. Langergraber, K. E. *et al. Proc. R. Soc. B* **278,** 408–416 (2011).
7. Boesch, C., Head, J. & Robbins, M. M. *J. Hum. Evol.* **56,** 560–569 (2009).
8. Mitani, J. C., Hasegawa, T., Gros-Louis, J., Marler, P. & Byrne, R. *Am. J. Primatol.* **27,** 233–243 (1992).



© 2011 Macmillan Publishers Limited. All rights reserved



# SHELVED

*For decades, the Armed Forces Institute of Pathology has been a leader in disease diagnosis. Now it is closing, and its legacy is in jeopardy.*

**BY ALISON MCCOOK**

The news took William Travis's breath away. The e-mail he had just opened revealed that the organization where he had worked for nearly 15 years — the Armed Forces Institute of Pathology (AFIP) in Washington DC — was to close. "I had cold chills," recalls Travis, now a thoracic pathologist at the Memorial Sloan-Kettering Cancer Center in New York City. "I was shocked. I thought, 'How could an institution like this be allowed to go away?' I had trouble breathing."

That was in the spring of 2005, just five months after Travis had left the institute. The storied medical centre had become a casualty in a wave of cost-cutting closures and consolidations at the US Department of Defense. And on 15 September, less than a year shy of its 150th birthday, the AFIP will shut its doors for good.

In its long history, the AFIP has become a stalwart of the international biomedical community. With its vast library of tissue samples and expertise at analysing tissues for the diagnosis of disease, it has been a valuable resource for researchers and clinicians alike. Every year, the AFIP received at least 50,000 requests for second opinions on difficult cases from external pathologists. The nearly 800 employees — including experts in many areas of human and animal pathology — made major or minor changes to roughly half of the cases they acted on.

The military has yet to decide whether academic scientists will continue to have access to the AFIP's unique tissue repository. The largest in the world, it holds 55 million glass slides, 31 million paraffin blocks and more than 500,000 wet tissue samples. Scientists had only just begun to apply modern molecular techniques, such as DNA and RNA sequencing, to the collection. "That repository is an international treasury. And that has got to be available to the community at large,"

K. KASMAUSKI/CORBIS

**With nearly 90 million samples, the AFIP's tissue repository is the world's largest.**



© 2011 Macmillan Publishers Limited. All rights reserved

# A diary of death & disease

The Armed Forces Institute of Pathology built up a worldwide reputation for its tissue repository and expertise in disease diagnosis.

**1862**
**THE BEGINNING**
Brigadier General William Hammond, the Army surgeon-general, establishes the Army Medical Museum to collect specimens from the American Civil War.



**1918**
**THE WAR**
The First World War brings nearly 10,000 specimens. The institution starts transforming into a diagnostic pathology facility.



**1944**
**STRANGE CASES**
Samples flood in during the Second World War, as soldiers return with mysterious diseases. The military establishes the Army Institute of Pathology to gather and learn from the cases.

NATIONAL MUSEUM OF HEALTH AND MEDICINE (LEFT); CORBIS

says John Madewell, a former AFIP radiologist now based at the M. D. Anderson Cancer Center in Houston, Texas. So even as scientists mourn the loss of the institution, many are wondering what will become of the fruits of its labours. "I am very concerned," says Madewell.

## HUMBLE BEGINNINGS

From the start, the AFIP had a mission that extended beyond military strategy. In the spring of 1862, three dried and varnished tissue specimens were placed on a shelf in the Washington DC office of Brigadier General William Hammond, the Army surgeon-general. It was the beginning of the Army Medical Museum (see 'A diary of death & disease'). Hammond wanted to collect and catalogue the specimens that had been accumulating from men fighting in the American Civil War. He wanted a collection that would "embrace all forms of injuries and diseases, so that eventually it would become a general pathological museum, accessible for study to all medical men who are prosecuting original inquiries".

The museum became increasingly important to the military during the two world wars, says Michael Rhode, an archivist and one of the few remaining AFIP employees, as soldiers came back with strange new diseases such as gas gangrene, in which bacteria produce tissue-killing toxins inside the body. The army decided it needed a central location to collect and learn from unique cases and in 1944, the museum established the Army Institute of Pathology. Five years later, the institute became the central laboratory of pathology for all branches of the armed forces, and adopted its current name.

Rhode says that "there was a strong tradition of involving civilians". It was in the military's interests to collect difficult cases from around the world, to learn more about the diseases troops could face. As the AFIP didn't initially charge for second opinions on difficult cases, the samples flooded in.

The institute's final home, on 16th Street in Washington DC, was built in the 1950s to accommodate the growing repository and the pathologists who wanted to work at the AFIP, drawn to the collection and the stream of unusual cases. Within the first five years in the new building, AFIP scientists conducted more than 200 investigations of misunderstood diseases.

AFIP pathologists also began producing the *Atlas of Tumor Pathology*, a set of frequently updated volumes that are considered "bibles" of the field, says Chris Kelly, a former spokesperson for the AFIP. "There is probably not a bookshelf in a pathology office that does not contain at least one of these."

But many clinicians would say that the AFIP had the most impact with its consultation service, in which resident experts provided a second opinion on cases submitted by external pathologists in which, for whatever reason, the original diagnosis was uncertain. For instance, says former AFIP pathologist Susan Abbondanzo, two high-grade lymphomas — lymphoblastic and Burkitt's lymphoma — can be difficult to distinguish using the typical tissue-staining method used by most pathologists. But each has a very different treatment, so an incorrect diagnosis could kill the patient. The AFIP could routinely diagnose these cases, relying on experts in conjunction with a molecular and immunohistochemical laboratory, which many small hospitals lack. "I'm not going to say every diagnosis is life-saving," says Abbondanzo. "But many of the things you see there are, potentially."

> *"That repository is an international treasury. And that has got to be available to the community."*

## A PATHOLOGIST'S GOLDMINE

When pathologists extol the value of the AFIP, the conversation always returns to the tissue repository — all told nearly 90 million tissue samples, including some of the most rare and difficult cases ever encountered in the history of the field. A general pathologist might see one or a few cases of a rare tissue abnormality during a career. The repository often contains many, enabling researchers to categorize the diseased tissues and standardize diagnosis.

Abbondanzo joined the AFIP's department of haematopathology in 1990, and the first research project she collaborated on was investigating whether the anti-seizure drug Dilantin caused tumours in lymph nodes. "It was a huge concern," she says. Abbondanzo had been a pathologist for six years, but had never seen a case of enlarged lymph nodes in people taking Dilantin. When she arrived at the AFIP, there were 25, which allowed Abbondanzo and her colleagues to determine that the vast majority of cases were benign, not cancerous — showing there was no association between the drug and the cancer[1].

And in the 1990s, Jeffery Taubenberger (then chair of the AFIP's department of molecular pathology) and his colleagues began applying molecular techniques to the repository's paraffin-embedded tissue blocks containing lung samples from soldiers killed by Spanish flu during the 1918 pandemic[2]. From these samples, among others, they were able to analyse the genome of the virus and investigate why it was so deadly. They traced its virulence to multiple genes, and found that it triggers a dramatic inflammatory response.

## OUT OF FAVOUR

Abbondanzo, who became the chair of the haematopathology department in 1994, began to suspect that the AFIP was in trouble during the first US war with Iraq in the early 1990s, when the Department of Defense began to take a closer look at its budget. Suddenly, research projects that would have been approved and encouraged in the past were being questioned. Abbondanzo and her colleagues had submitted a proposal to study follicular lymphoma, a type of cancer that is rare in children. The AFIP repository contained at least 20 childhood cases, providing an unprecedented opportunity to characterize the condition. But the board that approved proposals rejected the project, saying it had "no military relevancy", Abbondanzo says. During this same period, she attended a meeting at which someone from the department referred to the AFIP as an "obscure little agency".

Travis, too, was worried about the AFIP's future. In 2003, he and others asked pathologists from overseas to write letters urging the military to continue supporting the AFIP, and he worked with the office of Senator Edward Kennedy (Democrat, Massachusetts), a long-time supporter of the institute, which tried to get financial support from other agencies.



© 2011 Macmillan Publishers Limited. All rights reserved

**1949**
**GROWTH AND RECOGNITION**
The facility is renamed the Armed Forces Institute of Pathology (AFIP), and later moves to a new building in Washington DC.



**1997**
**DATA RESURRECTED**
A team led by Jeffery Taubenberger starts using lung samples from soldiers killed by the 1918 pandemic flu to analyse the viral genome.



**2001**
**GRIM TASK**
The AFIP mobilizes a team of more than 50 experts to identify the victims of the 11 September plane crashes at the Pentagon and Shanksville, Pennsylvania.



**2005**
**THE END NEARS**
The AFIP is officially 'disestablished' by the US Department of Defense. One year later, it receives its 3,000,000th case.

(FROM LEFT) AP; K. KASMAUSKI/SCIENCE FACTION/CORBIS; NATIONAL MUSEUM OF HEALTH AND MEDICINE

"I spent countless hours," Travis says. "All those efforts were not going anywhere. It was very clear I had to look out for myself and my family." When a position opened at Memorial Sloan-Kettering, he applied.

Despite rumours that some services might be cut, the closure announcement in 2005 came as a shock. "We learned of it when everyone else did," says Kelly. Rhode, who has worked at the AFIP since the 1980s, heard the news in an auditorium full of other staff. "I don't remember people talking much. It seemed like we kind of just spilled out and went back to our offices in silence."

The AFIP was on a list of base closures and changes designed to save more than US$30 billion over 20 years. But the institute's budget in fiscal year 2004 was only $93 million — slightly more than 0.02% of the defence department's budget request for 2004 of $380 billion. The military had welcomed civilian pathology cases for decades. Clearly, something had changed. A spokesperson for the Department of Defense said the decision to "disestablish" the AFIP was "based on capacity, military value and scenario development after analysis of military and non-military workload, services available in the civilian sector and cost savings".

What really caused the demise of the AFIP, speculates Abbondanzo, is that the vast majority of its work had limited direct benefit to the military. "It's been a long time coming," she says.

## LIFE AFTER DEATH

Not everything will disappear. A new entity, the Joint Pathology Center, has been created in Silver Spring, Maryland, to carry on the AFIP's military duties, including consulting on pathology cases for the military and other federal agencies.

The AFIP's museum, the National Museum of Health and Medicine, which includes such exhibits as the bullet that killed US President Abraham Lincoln, is being packed up and moved to Fort Detrick's Forest Glen Annex in Silver Spring — a "stressful" process, says Rhode. And Madewell has helped to move the radiological pathology training that the AFIP provided for the majority of US radiology residents to the American College of Radiology based in Reston, Virginia.

A 2007 report by the US Government Accountability Office concluded that the change in the AFIP's consultation and other services would have "minimal impact", because there are alternative sources of pathology expertise. In recent decades, some centres — such as the Mayo Clinic in Rochester, Minnesota, and Johns Hopkins University in Baltimore, Maryland — have developed strengths in particular fields of pathology, and have become sources for second opinions in difficult cases, says Colonel Vernon Armbrustmacher, one of the AFIP's former directors. "The world of pathology will survive without the AFIP," says Armbrustmacher. "I hate to say that, but it will."

Even so, these services could be more expensive than those of the AFIP, which charged from around $20 to just over $2,000 for consultations, depending on the procedure. Some pathologists will feel the hit more than others — especially those in developing countries, who frequently lack local services to diagnose emerging tropical diseases and can't afford to pay high fees. These were a minority of cases — perhaps between 5% and 10% of consultations, estimates Armbrustmacher. But "they were unbelievable cases", Travis recalls. "Really challenging, difficult problems."

> *"The world of pathology will survive without the AFIP. I hate to say that, but it will."*

And even when pathologists can afford second opinions from academic centres, what they receive may not be of the same calibre as what they would have got from the AFIP, says Timothy O'Leary, director of Clinical Science Research and Development at the US Veterans Health Administration in Washington DC, and a pathologist at the AFIP until 2004. "There are cases that are just not common. And at the AFIP, we saw a lot of those unusual cases, far more than you do at academic centres." It may only be "a few hundred, a thousand people each year" whose diagnoses might be significantly affected, he estimates. "But for those people, it can be a matter of life and death."

The fate of the tissue repository, which is now under the control of the Joint Pathology Center, remains unknown. It has been moved to two renovated buildings at Forest Glen Annex, one of which used to serve as the laundry facility for the AFIP and the Walter Reed Army Medical Center, which is also being relocated this year. Officials have asked the Institute of Medicine to recommend how best to use the repository, including who should have access to it, says the Joint Pathology Center's interim director Colonel Thomas Baker. Those recommendations are due in June 2012.

What's most important, many say, is that civilian pathologists continue to have access to the repository. New techniques in high-throughput genomics and proteomics could "now or soon" reveal even more clues about the deadly diseases preserved there, says Taubenberger, now at the National Institute of Allergy and Infectious Diseases in Bethesda, Maryland. The repository "is only of value to people who know what's in it", says Travis. If the top pathologists can't access the material, an untold amount of valuable knowledge will never be uncovered, he says.

Inside the AFIP, the mood is decidedly sombre as people watch the lights go out, says Rhode. "Every day, you see a little bit more furniture moving down the hallway." Little by little, more offices are emptied, cleaned, locked and fitted with 'do not enter' signs.

Florabel Mullick, the last director of the institute, declined to be interviewed for this article. But in the AFIP's final newsletter last winter, she lamented watching the institution become a "shell of its former self" and wrote that closing the AFIP "has been one of the most painful experiences of my life … How does one, after all, watch missions that have benefitted so many people fade away?"

"It is the passing of one of the greatest institutions in the entire history of medicine," agrees Travis. "And it is very painful to see that happen. But we all have to move on." ■

**Alison McCook** *is a Comment editor at* Nature.

1. Abbondanzo, S. L., Irey, N. S. & Frizzera, G. *Am. J. Surg. Pathol.* **19,** 675–686 (1995).
2. Taubenberger, J. K., Reid, A. H., Krafft, A. E., Bijwaard, K. E. & Fanning, T. G. *Science* **275,** 1793–1796 (1997).



© 2011 Macmillan Publishers Limited. All rights reserved

# COMMENT

**BRAIN** Memory-dampening drugs have more benefits than risks **p.275**



**HISTORY** When US quantum physics was quirky **p.278**

**TAXONOMY** Strip unwise *Homo* of *sapiens* rider until we mend our destructive ways **p.282**

**OBITUARY** John Marburger, science adviser to President Bush, remembered **p.284**

K. HARRISON



**Women and chemistry do mix: Carol V. Robinson (centre) with members of her research lab at the University of Oxford, UK.**

# In pursuit of female chemists

Chemistry needs new female role models and a less macho culture to appeal more to the next generation of young women, says **Carol V. Robinson**.

As the first female chemistry professor at both the University of Cambridge and the University of Oxford, which have a combined history in chemistry of about 800 years, I am often asked to comment on the poor retention of women chemists by UK universities. The decline from chemistry PhDs (46% women) to professorships (just 6%) is steeper than in other disciplines, including physics and engineering[1]. But numbers don't tell the whole story. Here I offer personal reflections from my career in chemistry about why women leave science.

The greatest attrition occurs during the transition from PhD to research. According to surveys done by the Royal Society of Chemistry in 2000 and 2008, this is when women become discouraged. They find laboratory research in chemistry too adversarial, the difficulties of combining career and family daunting and the lack of role models isolating[2,3].

My role model and that of my generation of women chemists is Marie Curie, whose 1911 Nobel Prize in Chemistry we celebrate this year to mark the International Year of Chemistry (she got her first Nobel, in physics, in 1903). Curie was often asked, especially by women, how she could combine family life with a scientific career. Her reply, "it has not been easy", resonates to this day. A century on from Curie's prize, as a mother of three, it is a question I, too, am often asked. One would hope for greater progress.

I was disappointed to read in a recent survey[4] that one-quarter of all 14-years-olds in the United Kingdom confuse Marie Curie with pop singer Mariah Carey. Perhaps we need new role models for today's women? ▶



© 2011 Macmillan Publishers Limited. All rights reserved

Curie's achievements might be too heroic and the sacrifices she made too great. If so, female chemists, including myself, can offer career narratives that defy the usual stereotypes and, hopefully, inspire others.

Unlike today's generation, I was inspired by Marie Curie's personal life story. Curie, a self-taught chemist, grew up in Warsaw; her father lost his job, forcing her to find work to help support her family. Similarly, my own full-time education ended at age 16 when my father lost his job. Consequently, I didn't follow the usual academic path, but after leaving school became a lab technician at Pfizer Corporation in Sandwich, UK.

Industry was a welcoming environment for me. I was passionately interested in mass spectrometry and began to keep notebooks of observations. My supervisors were impressed and encouraged me to attend night school. After seven years of part-time study and a year's full-time research for a master's degree, I qualified to study for a PhD at the University of Cambridge. At every stage, I felt I had to prove myself and worked hard. With limited PhD funding, I published five papers and wrote my thesis in less than two years.

Shortly after completing my PhD came the challenges of combining family life and career. I chose another unconventional route. I took an eight-year career break to cover the birth of my three children. I was warned that it was highly unlikely I would be able to return to science. I thought this was too high a price to pay for motherhood. Nowadays, when asked to talk to young women, I am often asked not to mention my career break, although I usually do. Sadly, it is not something that many institutions encourage.

## YOU CAN HAVE IT ALL

Is the choice between children — and a normal family life — and becoming a professional scientist really so stark? I asked Ada Yonath, the only living female Nobel laureate in chemistry, how she balanced the demands of family and science. She replied that it is possible as long as you have a passion for both. I agree — we have to change the notion that academic jobs cannot be family friendly.

Admittedly, after my eight-year absence, it was hard to find a position in science. I had three interviews before convincing a panel that I was committed (one interviewer remembered me positively from my student days). It also wasn't easy being a postdoc when my children were school-age, but I was immensely productive in the morning. From about 5 a.m. my thoughts were clear. I had no interruptions and no guilt. As my children got older, I began to attend international meetings, often taking the children with me, even into the lecture hall. These sorts of opportunities to mix family and professional life are unparalleled in many other professions.

My research may also have benefited from having a career break. On returning in 1992, well-meaning academics tried to persuade me to follow fashionable pathways in proteomics and, a few years later, in metabolomics. But becoming a principal investigator in my forties, much later than most, I was already several years behind the leading labs and not sufficiently excited by these trends. I needed to do something different.

*"Women don't need to adopt more 'male' character traits to succeed."*

"Nothing in life is to be feared, it is only to be understood," is my favourite quote from Curie, who, during her discovery of radioactivity, unknowingly exposed herself to great personal danger in her pursuit of the unknown. Fearlessness is one quality I have noticed among many successful chemists, men and women. Yonath did not take on easy projects and was regularly accused of being 'a dreamer'.

In my own lab, I pursued a path of putting macromolecular complexes into the gas phase of a mass spectrometer, not an obvious choice for the structural-biology questions I intended to ask. Well-respected scientists told me that the results would be meaningless. Happily, I chose not to follow too much of this advice. I might have been less brave had I been younger and more eager to please.

Chemistry has a reputation for being more cut-throat than biology or physics[5]. It has a macho culture in which getting to the finish line first is more important than how you get there. And women are assumed to be less competitive than men. There is some truth in this. Too often, female scientists shy away from responsible roles or don't have sufficient confidence or aspirations.

## CULTURE REINFORCEMENT

Certainly, we can't blame men alone for creating a male-dominated environment when women scientists are often too keen to be invisible, or worse, to become honorary males, by adopting aggressive communication styles, say, or eschewing any interest in appearance. Curie herself proudly claimed that she had no dresses save the black one she wore every day to the laboratory. When I started at the University of Cambridge, I was advised not to dress so well by a female colleague, who warned I would be confused with the secretarial staff. Another colleague confessed to hiding shopping bags for fear of being considered too vain. But by behaving or dressing as honorary men, we only reinforce the macho culture of chemistry.

Consider how off-putting that might be to the younger generation. Let's start with our own daughters. "Would you advise them to go into science?" was a question recently posed to a small group of top American women scientists[6]. "Only if she is made of steel," was the consensus. I don't agree. I would have been very happy for my only daughter, a mathematician, to have been a chemist. I am proud of a mother–daughter publication resulting from her holiday research in my lab. Women don't need to adopt more 'male' character traits to succeed.

Today, 11 of the 20 people in my lab are women. I didn't set out to make it this way. I think women are drawn to the group, believing that I am likely to understand the conflicting demands of family and career. I am particularly proud of the many successful young women with families whom I have mentored, but I don't feel that I am doing anything special. I simply treat everyone as I would like to be treated myself — supported and nurtured. Perhaps all we can do is to change attitudes one research group at a time.



LEFT TO RIGHT: HULTON-DEUTSCH COLLECTION/CORBIS; BETTMANN/CORBIS; KEYSTONE/GETTY IMAGES; M. PROBST/AP

**Four women have received the Nobel Prize in Chemistry: Marie Curie (1911), Irène Joliot-Curie (1935), Dorothy Hodgkin (1964) and Ada Yonath (2009).**



© 2011 Macmillan Publishers Limited. All rights reserved

We must also challenge the patronizing cultural stereotypes that deter women. Headlines such as "Oxford housewife wins Nobel" and "One giant leap for womankind — and Israel" (for Nobel laureates Dorothy Hodgkin and Yonath, respectively) are par for the course. "Mother wins prize" and "Woman wins top science prize", are both newspaper headlines reporting my more modest achievements. The Rosalind Franklin award in 2004, from the Royal Society, came with a significant sum described as "a lot of money to spend on shoes". I confess to a weakness for shoes. But I in fact set up a fund to support visiting female lecturers and to establish a mentoring scheme for female postgraduates at Cambridge.

Attitudes to women in science have improved, but there is still some way to go. Yonath remembers a time when a visitor to her lab commented that her research was "so good that he thought it likely carried out by a man". As a junior researcher, I was asked many times "who is the senior man you work for?". Today, female postgraduates note less explicit biases that can make them feel excluded: from the all-male photos in chemistry departments, to the timing of early evening seminars, and the ensuing discussions in the local pub. So, we must all do more to support female chemists: by speaking out at injustice or prejudice, instead of staying quiet, and by encouraging talented colleagues to aspire to more demanding roles.

Curie encouraged her elder daughter to follow in her footsteps. Irène Joliot-Curie received the second Nobel prize awarded to a woman in chemistry. As Marie Curie's granddaughter (herself a respected nuclear physicist) said: "We have to change the mentality that men are better scientists." Our daughters and granddaughters deserve nothing less. ■

**Carol V. Robinson** *is a Royal Society research professor, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, UK.*
*e-mail: carol.robinson@chem.ox.ac.uk*

1. Royal Society of Chemistry *Planning for Success: Good Practice in University Science Departments* (RSC, 2008).
2. Royal Society of Chemistry *Factors Affecting the Career Choices of Graduate Chemists* (RSC, 2000).
3. Royal Society of Chemistry *Change of Heart: Career Intentions and the Chemistry PhD* (RSC, 2008).
4. Royal Institution *7 Questions on Scientists* (BBC, 2011); available at go.nature.com/rbj63x
5. Newsome, J. L. *The Chemistry PhD: the Impact on Women's Retention* (RSC, 2008).
6. Kolata, G. Women atop their fields dissect the scientific life (*New York Times*, 6 June 2011).



ILLUSTRATION BY EDA AKALTUN

# Give memory-altering drugs a chance

The ethical challenges of memory-dampening drugs are likely to be manageable and the pay-offs considerable, says **Adam Kolber**.

Several studies suggest that memories can be pharmaceutically dampened. A few months ago, for instance, researchers showed that a drug called ZIP causes cocaine-addicted rats to forget the locations where they had regularly been receiving cocaine[1]. Other drugs, already tested in humans, may ease the emotional pain associated with memories of traumatic events. Indeed, the use of memory-altering drugs to treat addicts or victims of assaults, car accidents, natural disasters and terrorist attacks looks increasingly promising.

Many are alarmed by the prospect. As far back as 2003, the US President's Council on Bioethics issued a report that largely decried the use of such drugs[2]. Since then, journal articles and news stories have reiterated concerns that memory manipulators could interfere with the ability to lead true and honourable lives or could undermine a person's sense of identity.

The fears about pharmaceutical memory manipulation are overblown. Thoughtful regulation may some day be appropriate, but excessive hand-wringing now over the ethics of tampering with memory could stall research into preventing post-traumatic stress in millions of people. Delay could also hinder people who are already debilitated by harrowing memories from being offered the best hope yet of reclaiming their lives.

Various drugs are being investigated[3,4]. Propranolol, for instance — which is ▶



© 2011 Macmillan Publishers Limited. All rights reserved

We must also challenge the patronizing cultural stereotypes that deter women. Headlines such as "Oxford housewife wins Nobel" and "One giant leap for womankind — and Israel" (for Nobel laureates Dorothy Hodgkin and Yonath, respectively) are par for the course. "Mother wins prize" and "Woman wins top science prize", are both newspaper headlines reporting my more modest achievements. The Rosalind Franklin award in 2004, from the Royal Society, came with a significant sum described as "a lot of money to spend on shoes". I confess to a weakness for shoes. But I in fact set up a fund to support visiting female lecturers and to establish a mentoring scheme for female postgraduates at Cambridge.

Attitudes to women in science have improved, but there is still some way to go. Yonath remembers a time when a visitor to her lab commented that her research was "so good that he thought it likely carried out by a man". As a junior researcher, I was asked many times "who is the senior man you work for?". Today, female postgraduates note less explicit biases that can make them feel excluded: from the all-male photos in chemistry departments, to the timing of early evening seminars, and the ensuing discussions in the local pub. So, we must all do more to support female chemists: by speaking out at injustice or prejudice, instead of staying quiet, and by encouraging talented colleagues to aspire to more demanding roles.

Curie encouraged her elder daughter to follow in her footsteps. Irène Joliot-Curie received the second Nobel prize awarded to a woman in chemistry. As Marie Curie's granddaughter (herself a respected nuclear physicist) said: "We have to change the mentality that men are better scientists." Our daughters and granddaughters deserve nothing less. ■

**Carol V. Robinson** *is a Royal Society research professor, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, UK.*
*e-mail: carol.robinson@chem.ox.ac.uk*

1. Royal Society of Chemistry *Planning for Success: Good Practice in University Science Departments* (RSC, 2008).
2. Royal Society of Chemistry *Factors Affecting the Career Choices of Graduate Chemists* (RSC, 2000).
3. Royal Society of Chemistry *Change of Heart: Career Intentions and the Chemistry PhD* (RSC, 2008).
4. Royal Institution *7 Questions on Scientists* (BBC, 2011); available at go.nature.com/rbj63x
5. Newsome, J. L. *The Chemistry PhD: the Impact on Women's Retention* (RSC, 2008).
6. Kolata, G. Women atop their fields dissect the scientific life (*New York Times*, 6 June 2011).



ILLUSTRATION BY EDA AKALTUN

# Give memory-altering drugs a chance

The ethical challenges of memory-dampening drugs are likely to be manageable and the pay-offs considerable, says **Adam Kolber**.

Several studies suggest that memories can be pharmaceutically dampened. A few months ago, for instance, researchers showed that a drug called ZIP causes cocaine-addicted rats to forget the locations where they had regularly been receiving cocaine[1]. Other drugs, already tested in humans, may ease the emotional pain associated with memories of traumatic events. Indeed, the use of memory-altering drugs to treat addicts or victims of assaults, car accidents, natural disasters and terrorist attacks looks increasingly promising.

Many are alarmed by the prospect. As far back as 2003, the US President's Council on Bioethics issued a report that largely decried the use of such drugs[2]. Since then, journal articles and news stories have reiterated concerns that memory manipulators could interfere with the ability to lead true and honourable lives or could undermine a person's sense of identity.

The fears about pharmaceutical memory manipulation are overblown. Thoughtful regulation may some day be appropriate, but excessive hand-wringing now over the ethics of tampering with memory could stall research into preventing post-traumatic stress in millions of people. Delay could also hinder people who are already debilitated by harrowing memories from being offered the best hope yet of reclaiming their lives.

Various drugs are being investigated[3,4]. Propranolol, for instance — which is ▶



© 2011 Macmillan Publishers Limited. All rights reserved

already approved by the US Food and Drug Administration to treat hypertension — may dull the emotional pain associated with the recall of an upsetting experience by interfering with the release of stress hormones that otherwise strengthen memories. Preliminary studies suggest that propranolol can inhibit the formation of traumatic memories even when taken a few hours after a distressing event[5,6].

Given the close connection between memory and a sense of self, some bioethicists argue that instead of seeking a solution in a pill bottle, we should do the difficult but rewarding emotional work required to transform bad experiences into positive ones. They worry that giving people too much power to alter their life stories could ultimately weaken their sense of identity and make their lives less genuine[2].

These arguments are not persuasive. Some memories, such as those of rescue workers who clean up scenes of mass destruction, may have no redeeming value. Drugs may speed up the healing process more effectively than counselling, arguably making patients more true to themselves than they would be if a traumatic experience were to dominate their lives. As psychiatrist Peter Kramer described in his 1993 book *Listening to Prozac*, some patients report feeling like themselves for the first time after taking antidepressants.

Some have also argued that memory-dampening drugs may inure users to the pain of others. For example, soldiers might be less inhibited to kill if they think they can pharmaceutically dissipate the emotional impact of their actions[2]. Yet we permit soldiers to use remotely operated drones even though these devices are likely to make some soldiers more willing to kill. If we seek to minimize soldiers' physical wounds, we should also address their emotional ones.

Underlying these concerns is an unjustified aversion to pharmaceutical methods of managing trauma[7]. Almost every media report on memory-dampening drugs delves into the ethical issues they raise. Yet the ethical implications of non-pharmaceutical approaches to memory dampening are generally ignored. For example, a non-pharmaceutical method of helping subjects forget a learned association between a visual cue and an electric shock was described in *Nature* last year[8]. Although the study received widespread media coverage, ethicists largely ignored it.

Until researchers uncover the cost, efficacy and side effects both of pharmaceutical and of non-pharmaceutical approaches to memory dampening, one should not automatically be deemed inferior to the other.

That said, in some contexts there are grounds for regulating the use of drugs to manipulate memory — particularly drugs that alter factual recall.

Take a scenario recounted by a US doctor in 2007 (ref. 9). The doctor had biopsied a suspected cancer patient and sent a tissue sample to a pathologist while the woman was still in the operating room. Thinking she was completely sedated, the pathologist announced a bleak prognosis over the intercom. The patient, who had received only local anaesthesia, heard the news and began to shriek, "Oh my God. My kids!" An anaesthesiologist standing by quickly injected her with propofol, a sedative that causes some people to forget what happened a few minutes before they were injected. When the woman woke up, she had no memory of hearing her prognosis. Although the anaesthesiologist's action was well-meaning, the patient probably had the right to decide whether to have her mind manipulated[10].

Broad prohibitions, however, are unnecessary[7].In most countries, existing law would already prohibit particularly egregious uses. For example, getting a witness drunk to prevent him from testifying in court has long been deemed an obstruction of justice by US judges. Pharmaceutically erasing a witness's memory for the same purpose would already be considered a crime.

Lawmakers could easily add new restrictions. For example, in cases that could endanger people's lives, such as when the only witness to a violent crime seeks to forget traumatic memories, physicians could be required to contact police before prescribing memory-dampening drugs.

Fear that the potential fruits of research may be prohibited or heavily regulated could deter researchers from pursuing studies on memory manipulation or funders from supporting them. Instead, researchers should be encouraged to explore pharmaceutical and non-pharmaceutical methods of helping people cope with trauma. ■

**Adam Kolber** *is Professor of Law at Brooklyn Law School, New York 11201, USA, and editor of the Neuroethics & Law Blog.*
*e-mail: adam.kolber@brooklaw.edu*

1. Li. Y.-Q. *et al. J. Neurosci.* **31,** 5436–5446 (2011).
2. President's Council on Bioethics. *Beyond Therapy: Biotechnology and the Pursuit of Human Happiness* Ch. 5, 205–273 (2003).
3. Cukor, J., Spitalnick, J., Difede, J., Rizzo, A. & Rothbaum, B. O. *Clin. Psychol. Rev.* **29,** 715–726 (2009).
4. Marin, M.F., Hupbach, A., Maheu, F.S., Nader, K. & Lupien, S. J. *J. Clin. Endocrinol. Metab.* **96,** E1221–E1227 (2011).
5. Pitman, R. K. *et al. Biol. Psychiatry* **51,** 189–192 (2002).
6. Vaiva, G. *et al. Biol. Psychiatry* **54,** 947–949 (2003).
7. Kolber, A. J. *Vanderbilt Law Rev.* **59,** 1561–1626 (2006).
8. Schiller, D. *et al. Nature* **463,** 49–53 (2010).
9. Haig, S. The ethics of erasing a bad memory. *Time* (15 Oct. 2007); available at http://go.nature.com/jgov4l
10. Kolber, A. *Neuroethics* **1,** 145–148 (2008).


© 2011 Macmillan Publishers Limited. All rights reserved



COURTESY OF F. A. WOLF

**In the 1970s, Jack Sarfatti, Saul-Paul Sirag, Fred Alan Wolf and others in the Fundamental Fysiks Group opened up discussions of quantum mechanics.**

PHYSICS

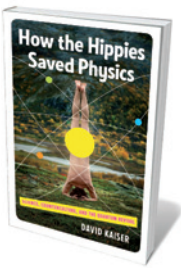
# Quantum outsiders

**Hugh Gusterson** enjoys a history of the quirky group that pursued quantum physics when it was unfashionable.

It is hard to write a book about quantum mechanics that is at once intellectually serious and a page-turner. But David Kaiser succeeds in his account of a neglected but important group of physicists who brought together quantum mechanics, Eastern religion, parapsychology and the hallucinogen LSD.

Kaiser, a historian of science at the Massachusetts Institute of Technology in Cambridge, seeks to understand why a set of mind-bending ideas developed in the 1920s and 1930s by European physicists — including Werner Heisenberg, Niels Bohr, Max Born and Wolfgang Pauli — were largely barred from discussion in mainstream US physics after the Second World War. At the heart of quantum mechanics lie these challenging ideas: that electrons behave like waves and particles; that one can know a particle's precise location or momentum but not both; that the act of observing a particle changes its behaviour; and that particles seem to communicate with one another across great distances — the phenomena of nonlocality and entanglement that Albert Einstein derided as "spooky action at a distance".

These radical concepts led to the field of quantum information science and the emergence of ultra-secure technologies, based on quantum encryption, for transferring money or electronic votes. Such concepts were also, in ways that Kaiser does not explore, influential beyond physics. Their critique of reductionism and emphasis on the effects of observers led to the rise of 'post-structuralist' thinking in the humanities. But during the cold war, the profound philosophical issues raised by quantum mechanics were seen by many in physics as an unnecessary distraction.




**How the Hippies Saved Physics: Science, Counterculture, and the Quantum Revival**
DAVID KAISER
*W. W. Norton: 2011. 372 pp. $26.95, £19.99*

Tracing shifts in university lecture notes and textbooks after the Second World War, Kaiser argues that the fundamental strangeness of quantum reality was drained out of academic physics by the "high-throughput pedagogy" that accompanied a sixfold increase in PhD student numbers. Pressure to produce gadgets to change the course of the cold war left little time for philosophizing.

This meant that some of the key work in quantum mechanics in the 1970s and 1980s was done by a motley crew of young physicists, who worked largely outside universities and published in obscure journals such as *Epistemological Letters* — "a hand-typed, mimeographed newsletter". They included Elizabeth Rauscher, Jack Sarfatti, Fred Alan Wolf, Saul-Paul Sirag, John Clauser and Fritjof Capra. The centre of their intellectual universe was the San Francisco Bay area. Many were associated with the Fundamental Fysiks Group, an open discussion group about quantum mechanics that started meeting in 1975 at the Lawrence Berkeley Laboratory in California.

Unable to secure professorships in the shrunken job market of the time, some of the group lived on the edge of destitution. The luckiest was Fritjof Capra, whose book *The Tao of Physics* (Wildwood House, 1975) was translated into 23 languages and brought him financial security. Other Fysiks members cobbled together eccentric funding sources: George Weissmann founded a company selling a Tibetan herbal remedy until the US Food and Drug Administration banned his product. And others secured funding through the Pentagon to see whether quantum mechanics might explain the apparent spoon-bending abilities of psychics such as Uri Geller and to investigate telepathy.

Three New Age guru–entrepreneurs were even more colourful. Werner Erhard, founder of the Est movement for self-transformation, sponsored opulent

NATURE.COM
David Kaiser on university funding:
go.nature.com/oflvik


© 2011 Macmillan Publishers Limited. All rights reserved

conferences on quantum mechanics at his San Francisco mansion. Members of the Fundamental Fysiks Group rubbed shoulders with Nobel laureates until the arrangement with Erhard fell apart in the face of growing concerns that Est was a cult. Michael Murphy, founder of the Esalen retreat on the California coast, sponsored workshops at which physicists alternated their hot-tub discussions of quantum mechanics with massages and, in some cases, LSD trips.

A third New Age activist and impresario, Ira Einhorn, somehow persuaded senior executives at Bell Telephone that the company should copy and mail the latest papers on quantum mechanics to some 300 people — an early postal version of an e-mail distribution list. This arrangement came to an abrupt end when the police discovered the decomposing body of Einhorn's girlfriend in a trunk in his apartment.

Interspersing entertaining anecdotes with explanations of concepts in quantum physics, Kaiser's book can be read on many levels. At its most challenging, it is an intellectual history of quantum mechanics. But it is also a yarn about an eccentric group of physicists who refused to be defeated by their marginalization within their own discipline. And, as social history, it offers a window onto one of the most colourful periods of twentieth-century US history.

The book makes important observations about the social dynamics of physics in the United States during the cold war. Kaiser argues that, even as military patronage pumped massive financial resources into physics, the discipline's horizons shrank. University physicists disdained the philosophical questions that had enlivened pre-war European physics and developed a narrowly instrumentalist pedagogy that sometimes became a straitjacket.

Kaiser describes some students holding secret meetings to discuss quantum mechanics behind their advisers' backs, having been warned that "thinking seriously about foundations was a waste of time and a detriment to one's career". He also notes that the editor of *Physical Review* banned articles discussing interpretations of quantum mechanics; a brilliant experiment on Bell's theorem by John Clauser was scarcely cited because of the prevailing orthodoxy. Clauser was told that his experiment was not "real physics", and he had a terrible time on the job market. By contrast, the heroes of Kaiser's story "strove to expand the physics profession's collective mental space". This is an equally apt description of Kaiser's approach in this illuminating book. ■

**Hugh Gusterson** *is professor of anthropology and sociology at George Mason University, Fairfax, Virginia 22030, USA. His latest book is* People of the Bomb. *e-mail: hgusters@gmu.edu*

EVOLUTION

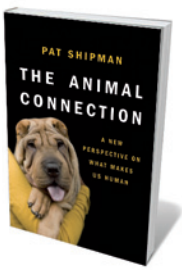
# The long trek to domesticated bliss

## Our fondness for fauna comes from an evolved human need to nurture, argues **Juliet Clutton-Brock**.

In *The Animal Connection*, palaeoanthropologist Pat Shipman describes how humans have been connected with animals for the past 2.6 million years, and how this relationship distinguishes us from other animals. She argues that the evolution of human behaviour was driven by this connection in three stages: toolmaking, language and domestication of animals and plants. But Shipman's book is more a review of what has long been known than the "New Perspective" promised by its subtitle.

Shipman starts by outlining the characteristics that separate the fossil hominin *Ardipithecus ramidus* — found in Ethiopia and dated to around 4.2 million years ago — from other primates. Continuing on through the australopithecines, she describes how, from 2.6 million years ago, "Hominids transformed rocks into stone tools and stone tools transformed hominids from bipedal apes that are [*sic*] basically herbivorous into predators". This was the beginning of the animal connection, accompanying the evolution of *Homo erectus* and its expansion out of Africa.

Shipman goes on to discuss the evolution of human language, the origin of symbolism and the theory of mind. She proposes that language followed from the need for humans to communicate animal-related information in their quest for food. Language "allowed humans to talk about animals and in time, with animals", she says.




**The Animal Connection: A New Perspective on What Makes us Human**
PAT SHIPMAN
*W. W. Norton: 2011. 336 pp. $26.95, £20*

Towards the end of the Pleistocene, around 32,000 years ago, came the first domestic dogs. The cultivation of plants and the domestication of livestock in many parts of the world followed in the early Holocene. Shipman describes what is known about this process, including arguments around the outdated view that the primary motive for domestication was the provision of meat. She also explores 'self-domestication' — the theory that, from



W. R. BILENDUKE/GETTY

**Domesticated 32,000 years ago, dogs remain central to many societies, including the Inuit.**



© 2011 Macmillan Publishers Limited. All rights reserved

conferences on quantum mechanics at his San Francisco mansion. Members of the Fundamental Fysiks Group rubbed shoulders with Nobel laureates until the arrangement with Erhard fell apart in the face of growing concerns that Est was a cult. Michael Murphy, founder of the Esalen retreat on the California coast, sponsored workshops at which physicists alternated their hot-tub discussions of quantum mechanics with massages and, in some cases, LSD trips.

A third New Age activist and impresario, Ira Einhorn, somehow persuaded senior executives at Bell Telephone that the company should copy and mail the latest papers on quantum mechanics to some 300 people — an early postal version of an e-mail distribution list. This arrangement came to an abrupt end when the police discovered the decomposing body of Einhorn's girlfriend in a trunk in his apartment.

Interspersing entertaining anecdotes with explanations of concepts in quantum physics, Kaiser's book can be read on many levels. At its most challenging, it is an intellectual history of quantum mechanics. But it is also a yarn about an eccentric group of physicists who refused to be defeated by their marginalization within their own discipline. And, as social history, it offers a window onto one of the most colourful periods of twentieth-century US history.

The book makes important observations about the social dynamics of physics in the United States during the cold war. Kaiser argues that, even as military patronage pumped massive financial resources into physics, the discipline's horizons shrank. University physicists disdained the philosophical questions that had enlivened pre-war European physics and developed a narrowly instrumentalist pedagogy that sometimes became a straitjacket.

Kaiser describes some students holding secret meetings to discuss quantum mechanics behind their advisers' backs, having been warned that "thinking seriously about foundations was a waste of time and a detriment to one's career". He also notes that the editor of *Physical Review* banned articles discussing interpretations of quantum mechanics; a brilliant experiment on Bell's theorem by John Clauser was scarcely cited because of the prevailing orthodoxy. Clauser was told that his experiment was not "real physics", and he had a terrible time on the job market. By contrast, the heroes of Kaiser's story "strove to expand the physics profession's collective mental space". This is an equally apt description of Kaiser's approach in this illuminating book. ■

**Hugh Gusterson** *is professor of anthropology and sociology at George Mason University, Fairfax, Virginia 22030, USA. His latest book is* People of the Bomb. *e-mail: hgusters@gmu.edu*

EVOLUTION

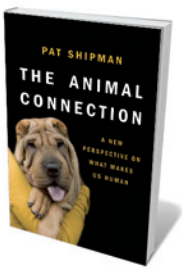
# The long trek to domesticated bliss

## Our fondness for fauna comes from an evolved human need to nurture, argues **Juliet Clutton-Brock**.

In *The Animal Connection*, palaeoanthropologist Pat Shipman describes how humans have been connected with animals for the past 2.6 million years, and how this relationship distinguishes us from other animals. She argues that the evolution of human behaviour was driven by this connection in three stages: toolmaking, language and domestication of animals and plants. But Shipman's book is more a review of what has long been known than the "New Perspective" promised by its subtitle.

Shipman starts by outlining the characteristics that separate the fossil hominin *Ardipithecus ramidus* — found in Ethiopia and dated to around 4.2 million years ago — from other primates. Continuing on through the australopithecines, she describes how, from 2.6 million years ago, "Hominids transformed rocks into stone tools and stone tools transformed hominids from bipedal apes that are [*sic*] basically herbivorous into predators". This was the beginning of the animal connection, accompanying the evolution of *Homo erectus* and its expansion out of Africa.

Shipman goes on to discuss the evolution of human language, the origin of symbolism and the theory of mind. She proposes that language followed from the need for humans to communicate animal-related information in their quest for food. Language "allowed humans to talk about animals and in time, with animals", she says.




**The Animal Connection: A New Perspective on What Makes us Human**
PAT SHIPMAN
*W. W. Norton: 2011. 336 pp. $26.95, £20*

Towards the end of the Pleistocene, around 32,000 years ago, came the first domestic dogs. The cultivation of plants and the domestication of livestock in many parts of the world followed in the early Holocene. Shipman describes what is known about this process, including arguments around the outdated view that the primary motive for domestication was the provision of meat. She also explores 'self-domestication' — the theory that, from



W. R. BILENDUKE/GETTY

**Domesticated 32,000 years ago, dogs remain central to many societies, including the Inuit.**



© 2011 Macmillan Publishers Limited. All rights reserved

scavenging around human settlements, wild animals gradually became habituated, were selected for tameness, and over generations became domesticated. Shipman outlines the distinction of true domestication from taming as "the permanence of the change in the wild animal". Yet she misses what I believe is the essence of why humans relate so readily to animals, which is supported by Sarah Blaffer Hrdy in her book on the evolution of hominins as cooperative breeders, *Mothers and Others* (Harvard University Press, 2009).

Humans, Blaffer Hrdy says, alone among the great apes, readily nurture each other's children. Without this help, few children in hunter-gatherer societies would survive to adulthood. There are many examples of hunter-gatherers extending the shared care of their infants to the adoption of young animals. So the human desire to enfold other species within our societies may be explained as having evolved from the combined instincts for nurture and domination.

I believe that the inborn human desire to nurture children and animals was followed by the domestication of dogs, and all the livestock animals whose social behaviour allowed it, in hunter-gatherer societies that were under pressure in the early Holocene from population growth and climate change. Shipman describes the conventional theory that domestication follows from the selection by humans of favoured attributes. I see the process as more complicated and in two parts: biological and cultural.

It is now accepted that some wild animals have cultures, that is, the inheritance of learned behaviour. With taming, an animal is brought into a protected place where it learns a new set of social relationships, as well as new feeding and reproductive strategies. Biological domestication is complete only when this 'culture' becomes heritable.

Shipman ends with the conviction that the ancient, innate connection between humans and animals is grossly underestimated in today's urban landscape. I see little evidence for this. Despite the inexorable spread of megacities and factory farms, the connection with both domestic and wild animals still occupies the minds and lives of innumerable people around the world. ■

**Juliet Clutton-Brock** *is a research associate at the Natural History Museum, London. Her forthcoming book is* Animals as Domesticates: A World View Through History.
*e-mail: juliet.cb@btinternet.com*

ASTRONOMY

# Supernova century

**Jay Pasachoff** relishes a novel that brings to life the scientific stars of the 1600s.

Somewhat eclipsed in popular culture by his great Italian contemporary Galileo Galilei, Johannes Kepler is finally achieving the iconic status he has long held in science. In 2009, the seventeenth-century German astronomer was the subject of an opera by US composer Philip Glass (*Nature* **462,** 724; 2009). Now, Stuart Clark's novel — the first in a trilogy about famous astronomers — puts fictional flesh on the bones of Kepler's life and times to enjoyable effect.

Only 60 years or so after Copernicus provided the idea of the heliocentric Universe, Kepler worked out the orbits of the planets. The story told in *The Sky's Dark Labyrinth,* which takes its name from a phrase in Galileo's 1623 book *Il Saggiatore* (*The Assayer*), is well known. Kepler assisted aristocratic court astronomer Tycho Brahe in Prague, taking over Tycho's precise observations of Mars and its changing position in the sky. From these studies, Kepler deduced his three laws of planetary motion, the first two of which he published in *Astronomia Nova* (*The New Astronomy*) in 1609, the same year that Galileo first pointed his telescope skywards.

Clark depicts the clash of two strong personalities: the haughty Tycho, and Kepler, whose confidence in his own mathematical abilities never wavered. He draws, too,




**Astronomer Johannes Kepler in his final decade.**




**The Sky's Dark Labyrinth: A Novel**
STUART CLARK
*Polygon: 2011. 272 pp. £12.99*

on the interaction of each with Rudolph II, the Holy Roman Emperor, who became Kepler's patron and whose name is commemorated in the Rudolphine Tables (1627). These were produced by Kepler using the laws of planetary orbits, which he derived on the basis of Tycho's observations. Working with these 'planetary tables', Kepler accurately predicted that transits of Mercury and of Venus would occur in 1631. The first of these was observed in Europe, but he did not realize that there would soon be a second transit of Venus, in 1639. (The next transit of Venus will be visible from Earth on 5–6 June 2012, depending on the observer's location; the following one is not until 2117.)

The author paints the conflicts between Lutherans and Catholics that drove the Lutheran Kepler from Graz to Prague, and that helped govern how Pope Urban VIII treated Galileo. Clark describes the blood that literally flowed during the internecine warfare between Rudolph II and his brother, Matthias, as the latter's troops attacked while Kepler and his family cowered in the city — evoking parallels with battles today.

My wife and I have made several astronomy-related pilgrimages: to Prague to see the plaque over Kepler's lodgings and his joint statue with Tycho; to dine at the Golden Griffin where Tycho lodged for a time, now a restaurant and hotel; to visit a monument (we found it defaced) to Kepler in Regensburg, Bavaria, possibly near where his bones were originally buried until they were lost; and to see the house in Regensburg where Kepler died in 1630, now a museum. I was also able to help the Houghton Library at Harvard University in Cambridge, Massachusetts, acquire the only known copy of Kepler's 1603 almanac.

As Clark emphasizes, 1603 was thought to be particularly auspicious at the time. Rudolph II is quoted as saying, "Eight hundred years earlier, Charlemagne founded Europe; eight hundred years before him, Christ was born." Soon thereafter, Kepler saw a supernova — the last seen in our

INTERFOTO/ALAMY



© 2011 Macmillan Publishers Limited. All rights reserved

scavenging around human settlements, wild animals gradually became habituated, were selected for tameness, and over generations became domesticated. Shipman outlines the distinction of true domestication from taming as "the permanence of the change in the wild animal". Yet she misses what I believe is the essence of why humans relate so readily to animals, which is supported by Sarah Blaffer Hrdy in her book on the evolution of hominins as cooperative breeders, *Mothers and Others* (Harvard University Press, 2009).

Humans, Blaffer Hrdy says, alone among the great apes, readily nurture each other's children. Without this help, few children in hunter-gatherer societies would survive to adulthood. There are many examples of hunter-gatherers extending the shared care of their infants to the adoption of young animals. So the human desire to enfold other species within our societies may be explained as having evolved from the combined instincts for nurture and domination.

I believe that the inborn human desire to nurture children and animals was followed by the domestication of dogs, and all the livestock animals whose social behaviour allowed it, in hunter-gatherer societies that were under pressure in the early Holocene from population growth and climate change. Shipman describes the conventional theory that domestication follows from the selection by humans of favoured attributes. I see the process as more complicated and in two parts: biological and cultural.

It is now accepted that some wild animals have cultures, that is, the inheritance of learned behaviour. With taming, an animal is brought into a protected place where it learns a new set of social relationships, as well as new feeding and reproductive strategies. Biological domestication is complete only when this 'culture' becomes heritable.

Shipman ends with the conviction that the ancient, innate connection between humans and animals is grossly underestimated in today's urban landscape. I see little evidence for this. Despite the inexorable spread of megacities and factory farms, the connection with both domestic and wild animals still occupies the minds and lives of innumerable people around the world. ■

**Juliet Clutton-Brock** *is a research associate at the Natural History Museum, London. Her forthcoming book is* Animals as Domesticates: A World View Through History.
*e-mail: juliet.cb@btinternet.com*

ASTRONOMY

# Supernova century

**Jay Pasachoff** relishes a novel that brings to life the scientific stars of the 1600s.

Somewhat eclipsed in popular culture by his great Italian contemporary Galileo Galilei, Johannes Kepler is finally achieving the iconic status he has long held in science. In 2009, the seventeenth-century German astronomer was the subject of an opera by US composer Philip Glass (*Nature* **462,** 724; 2009). Now, Stuart Clark's novel — the first in a trilogy about famous astronomers — puts fictional flesh on the bones of Kepler's life and times to enjoyable effect.

Only 60 years or so after Copernicus provided the idea of the heliocentric Universe, Kepler worked out the orbits of the planets. The story told in *The Sky's Dark Labyrinth,* which takes its name from a phrase in Galileo's 1623 book *Il Saggiatore* (*The Assayer*), is well known. Kepler assisted aristocratic court astronomer Tycho Brahe in Prague, taking over Tycho's precise observations of Mars and its changing position in the sky. From these studies, Kepler deduced his three laws of planetary motion, the first two of which he published in *Astronomia Nova* (*The New Astronomy*) in 1609, the same year that Galileo first pointed his telescope skywards.

Clark depicts the clash of two strong personalities: the haughty Tycho, and Kepler, whose confidence in his own mathematical abilities never wavered. He draws, too,




**Astronomer Johannes Kepler in his final decade.**




**The Sky's Dark Labyrinth: A Novel**
STUART CLARK
*Polygon: 2011. 272 pp. £12.99*

on the interaction of each with Rudolph II, the Holy Roman Emperor, who became Kepler's patron and whose name is commemorated in the Rudolphine Tables (1627). These were produced by Kepler using the laws of planetary orbits, which he derived on the basis of Tycho's observations. Working with these 'planetary tables', Kepler accurately predicted that transits of Mercury and of Venus would occur in 1631. The first of these was observed in Europe, but he did not realize that there would soon be a second transit of Venus, in 1639. (The next transit of Venus will be visible from Earth on 5–6 June 2012, depending on the observer's location; the following one is not until 2117.)

The author paints the conflicts between Lutherans and Catholics that drove the Lutheran Kepler from Graz to Prague, and that helped govern how Pope Urban VIII treated Galileo. Clark describes the blood that literally flowed during the internecine warfare between Rudolph II and his brother, Matthias, as the latter's troops attacked while Kepler and his family cowered in the city — evoking parallels with battles today.

My wife and I have made several astronomy-related pilgrimages: to Prague to see the plaque over Kepler's lodgings and his joint statue with Tycho; to dine at the Golden Griffin where Tycho lodged for a time, now a restaurant and hotel; to visit a monument (we found it defaced) to Kepler in Regensburg, Bavaria, possibly near where his bones were originally buried until they were lost; and to see the house in Regensburg where Kepler died in 1630, now a museum. I was also able to help the Houghton Library at Harvard University in Cambridge, Massachusetts, acquire the only known copy of Kepler's 1603 almanac.

As Clark emphasizes, 1603 was thought to be particularly auspicious at the time. Rudolph II is quoted as saying, "Eight hundred years earlier, Charlemagne founded Europe; eight hundred years before him, Christ was born." Soon thereafter, Kepler saw a supernova — the last seen in our

INTERFOTO/ALAMY



© 2011 Macmillan Publishers Limited. All rights reserved

Galaxy and now named after him.

I could all but smell the streets and markets of seventeenth-century Prague in this novel. In one memorable passage, Clark describes Kepler taking the path to the castle, across the Stone Bridge, where in 1611 he wrote his little book on the snowflake, recently republished to celebrate its 400th anniversary.

Clark also brings to life interesting minor characters and conjures up Kepler's eventful family history, including the joys of parenthood, his difficult mother (whom he had to clear of suspicions of witchcraft), and the tragic deaths of his children and first wife, Barbara. Thoughtfully crafted dialogues reveal the tension between Kepler's rationalism and the 'magical' beliefs of others.

Interspersed among the chapters about Kepler are several about Galileo's time in Padua, Florence and Rome. In these, Kepler endorses the veracity of Galileo's reports of seeing new 'stars' around Jupiter through his newfangled optical tube. And Galileo's lack of response to a letter from Kepler is explained as a result of religious rivalry: the Catholic church in Rome feared that lapsed-Lutheran Kepler would side with their Protestant enemies.

***"The fun of reading plausible words from the mouths of Kepler and Galileo overwhelms objections to invented conversations."***

Today, in an age when Vatican astronomers have telescopes in Arizona and host summer schools on cosmology at the Pope's summer residence at Castel Gandolfo near Rome, it is hard to evoke the mood of the early seventeenth century. Clark manages it.

Just as Shakespeare's inventions about the lives of kings give us too much pleasure to resist his playing fast and loose with history, the fun of reading plausible words from the mouths of Kepler, Galileo and their contemporaries overwhelms petty objections to invented conversations. I am less fond of Clark's minor chronological adjustments and his invented character, Cardinal Pippe.

*The Sky's Dark Labyrinth* deserves a broad readership. I look forward to Clark's forthcoming novel on Isaac Newton and Edmond Halley, and the final part of the trilogy, on Albert Einstein, Edwin Hubble and Georges Lemaître. ■

**Jay M. Pasachoff** *is professor of astronomy at Williams College, Williamstown, Massachusetts, USA.*
*e-mail: jay.m.pasachoff@williams.edu*

# Q&A Daphne Sheldrick
# Elephant rescuer

*Daphne Sheldrick was the first person to rear baby elephants successfully by hand, and has worked with animals for 50 years in Kenya. As she stars in an IMAX film chronicling her efforts, she describes her experience of conservation and animal husbandry.*

**How did you get involved with elephants?**

I grew up on a highland farm in Kenya. We had a lot of wild orphan animals because in those days wild animals were everywhere. I married into the wildlife service and lived in Tsavo National Park, where wild orphans started coming in — buffalo, rhinos and elephants. The elephants were a huge challenge: nobody had managed to raise a newborn calf. My husband thought it was impossible. Eventually I managed to keep one alive for six months, having lost many others. We knew we were on the right track.

**Born to Be Wild 3D**
*On release now at IMAX cinemas.*

**Why are elephants so challenging to rear?**

The milk formula is very special. So is the husbandry. It took 28 years of trial and error to get both about right. We knew that elephants' milk was high in fat. We added extra cream and butter to cows' milk, but then we learned that the elephants lived a lot longer on skimmed milk. So we scouted around the shops to try to find infant human formula that contained vegetable fat instead — the nearest thing to the fat in elephants' milk is coconut fat. With this mix, we have now managed to raise 130 elephants, which are still living today, some with their own young.

**How do these animals become orphaned?**

Elephants are fragile in infancy. In a perfect world they would not have any enemies, but the growing human population is taking up more and more land that was once used for elephant migration routes. An elephant does not know the difference between grass and maize [corn] and can eat a human's livelihood in an afternoon. And the demand in the Far East for ivory has grown tremendously in the past few years. The only hope for elephants is to ban all trade in ivory.

**Have you shared your methods?**

We have tried to. We are contacted by people from all over Africa, but raising elephants is a long-term thing. They are with you for up to 10 years, sometimes longer. They need 24-hour care and feeding every 3 hours, day and night. They are milk-dependent for 3 years. You have got to have the staying power to see the project through. I've been working with elephants for 50 years — that's most of my life.

**NATURE.COM**
More on the surge in the ivory trade:
go.nature.com/k9sgfk



DAVID SHELDRICK WILDLIFE TRUST (WWW.SHELDRICKWILDLIFETRUST.ORG)

**Daphne Sheldrick with one of her early charges.**

**How do you return hand-raised elephants to the wild?**

They are one of the easiest species to return; it just takes a long time. It isn't a question of getting an animal to two or three years old and tipping it out in the bush. They go to the rehabilitation stations in Tsavo with their human family until they have made friends in the wild herds. They are introduced to the wild herds by other now-integrated orphans. No elephant we have raised has been unable to go wild in the fullness of time.

**What does the future hold for elephants in Africa?**

The elephant population in Africa will fall to a much lower level as the human population rises. But the absence of elephants has a big impact. They create the water holes and the trails that lead to them; they turn thick unproductive scrubland into productive grassland; they walk enormous distances and distribute seeds in their dung. Without the elephant you lose lots of other animals.

**What do you hope people will take away from *Born to be Wild 3D*?**

Everyone that sees it cannot fail to go home having learned a lot about elephants — how caring and nurturing they are and how magnificent they are to one another. Elephants can provide lots of lessons for humans. ■ SEE CORRESPONDENCE P.282

INTERVIEW BY DANIEL CRESSEY



© 2011 Macmillan Publishers Limited. All rights reserved

Galaxy and now named after him.

I could all but smell the streets and markets of seventeenth-century Prague in this novel. In one memorable passage, Clark describes Kepler taking the path to the castle, across the Stone Bridge, where in 1611 he wrote his little book on the snowflake, recently republished to celebrate its 400th anniversary.

Clark also brings to life interesting minor characters and conjures up Kepler's eventful family history, including the joys of parenthood, his difficult mother (whom he had to clear of suspicions of witchcraft), and the tragic deaths of his children and first wife, Barbara. Thoughtfully crafted dialogues reveal the tension between Kepler's rationalism and the 'magical' beliefs of others.

Interspersed among the chapters about Kepler are several about Galileo's time in Padua, Florence and Rome. In these, Kepler endorses the veracity of Galileo's reports of seeing new 'stars' around Jupiter through his newfangled optical tube. And Galileo's lack of response to a letter from Kepler is explained as a result of religious rivalry: the Catholic church in Rome feared that lapsed-Lutheran Kepler would side with their Protestant enemies.

***"The fun of reading plausible words from the mouths of Kepler and Galileo overwhelms objections to invented conversations."***

Today, in an age when Vatican astronomers have telescopes in Arizona and host summer schools on cosmology at the Pope's summer residence at Castel Gandolfo near Rome, it is hard to evoke the mood of the early seventeenth century. Clark manages it.

Just as Shakespeare's inventions about the lives of kings give us too much pleasure to resist his playing fast and loose with history, the fun of reading plausible words from the mouths of Kepler, Galileo and their contemporaries overwhelms petty objections to invented conversations. I am less fond of Clark's minor chronological adjustments and his invented character, Cardinal Pippe.

*The Sky's Dark Labyrinth* deserves a broad readership. I look forward to Clark's forthcoming novel on Isaac Newton and Edmond Halley, and the final part of the trilogy, on Albert Einstein, Edwin Hubble and Georges Lemaître. ■

**Jay M. Pasachoff** *is professor of astronomy at Williams College, Williamstown, Massachusetts, USA.*
*e-mail: jay.m.pasachoff@williams.edu*

# Q&A Daphne Sheldrick
# Elephant rescuer

*Daphne Sheldrick was the first person to rear baby elephants successfully by hand, and has worked with animals for 50 years in Kenya. As she stars in an IMAX film chronicling her efforts, she describes her experience of conservation and animal husbandry.*

**How did you get involved with elephants?**

I grew up on a highland farm in Kenya. We had a lot of wild orphan animals because in those days wild animals were everywhere. I married into the wildlife service and lived in Tsavo National Park, where wild orphans started coming in — buffalo, rhinos and elephants. The elephants were a huge challenge: nobody had managed to raise a newborn calf. My husband thought it was impossible. Eventually I managed to keep one alive for six months, having lost many others. We knew we were on the right track.

**Born to Be Wild 3D**
*On release now at IMAX cinemas.*

**Why are elephants so challenging to rear?**

The milk formula is very special. So is the husbandry. It took 28 years of trial and error to get both about right. We knew that elephants' milk was high in fat. We added extra cream and butter to cows' milk, but then we learned that the elephants lived a lot longer on skimmed milk. So we scouted around the shops to try to find infant human formula that contained vegetable fat instead — the nearest thing to the fat in elephants' milk is coconut fat. With this mix, we have now managed to raise 130 elephants, which are still living today, some with their own young.

**How do these animals become orphaned?**

Elephants are fragile in infancy. In a perfect world they would not have any enemies, but the growing human population is taking up more and more land that was once used for elephant migration routes. An elephant does not know the difference between grass and maize [corn] and can eat a human's livelihood in an afternoon. And the demand in the Far East for ivory has grown tremendously in the past few years. The only hope for elephants is to ban all trade in ivory.

**Have you shared your methods?**

We have tried to. We are contacted by people from all over Africa, but raising elephants is a long-term thing. They are with you for up to 10 years, sometimes longer. They need 24-hour care and feeding every 3 hours, day and night. They are milk-dependent for 3 years. You have got to have the staying power to see the project through. I've been working with elephants for 50 years — that's most of my life.

**NATURE.COM**
More on the surge in the ivory trade:
go.nature.com/k9sgfk



DAVID SHELDRICK WILDLIFE TRUST (WWW.SHELDRICKWILDLIFETRUST.ORG)

**Daphne Sheldrick with one of her early charges.**

**How do you return hand-raised elephants to the wild?**

They are one of the easiest species to return; it just takes a long time. It isn't a question of getting an animal to two or three years old and tipping it out in the bush. They go to the rehabilitation stations in Tsavo with their human family until they have made friends in the wild herds. They are introduced to the wild herds by other now-integrated orphans. No elephant we have raised has been unable to go wild in the fullness of time.

**What does the future hold for elephants in Africa?**

The elephant population in Africa will fall to a much lower level as the human population rises. But the absence of elephants has a big impact. They create the water holes and the trails that lead to them; they turn thick unproductive scrubland into productive grassland; they walk enormous distances and distribute seeds in their dung. Without the elephant you lose lots of other animals.

**What do you hope people will take away from *Born to be Wild 3D*?**

Everyone that sees it cannot fail to go home having learned a lot about elephants — how caring and nurturing they are and how magnificent they are to one another. Elephants can provide lots of lessons for humans. ■ SEE CORRESPONDENCE P.282

INTERVIEW BY DANIEL CRESSEY



© 2011 Macmillan Publishers Limited. All rights reserved

# CORRESPONDENCE

## Fishery reform: many stocks secure

We contest the widely held view that all European fish stocks and fisheries are in a dire state (see, for example, *Nature* **475,** 7; 2011).

With the exception of cod, which remains heavily overfished, key stocks such as haddock, whiting, plaice and sole are all exploited close to the maximum sustainable yield.

We examined data for the 40 fish stocks assessed by the International Council for the Exploration of the Sea (ICES) that have reference points for maximum sustainable yields. These account for more than 90% of total allowable catches by European Union (EU) fishing fleets in the ICES areas.

The average fishing mortality for these stocks has declined in the past 10 years and their average spawning stock biomass has increased. In 2010, 17 of them (43%) were fished at or below the level that delivers maximum sustainable yields — a substantial improvement on 12.5% in 2000. Fishing mortality has fallen since 2000 from 2.1 times the maximum sustainable yield to 1.3 times in 2010.

Much remains to be done, but there are clear indications that actions implemented under the EU's Common Fisheries Policy (CFP) have been effective in bringing some stocks towards the target that will allow maximum sustainable yields. The CFP is under reform not because it has failed to achieve its targets, but because European Common Policies are reviewed every ten years (the last was in 2002).

After several decades of overexploitation, achieving this objective for all stocks will take time, but it is well under way. For some seas — the Mediterranean, for example — where the situation for most fish stocks is critical, we do not yet know whether management measures are delivering their intended objectives.

**Massimiliano Cardinale** *on behalf of 9 co-signatories*, Swedish University of Agricultural Sciences, Lysekil, Sweden. massimiliano.cardinale@slu.se*

**See http://dx.doi.org/10.1038/476282a for a full list of signatories.*




## New name needed for unwise *Homo*?

*Homo sapiens* was the name Carl Linnaeus assigned to our species in 1758, when humanity may indeed have seemed 'wise' relative to others. Today, this name is up for debate, given our questionable ability to control the potential disasters we have unleashed.

An animal that imperils its own future and that of most other life forms and ecosystems does not merit a single '*sapiens*', let alone the two we now bear.

We should be formally renamed to more accurately describe a species that is: exterminating thousands of others; releasing carbon, nitrogen and phosphorus in amounts exceeding Earth's natural cycles; devoting 50 times more resources to making weapons than to sustaining the food supply; destroying forests; degrading soil; polluting water; pillaging the oceans; and damaging the atmosphere on a planetary scale.

Changing our species' name might risk infringing some of the hallowed rules of nomenclature, but it would send an important signal about our present collective behaviour. Our present name *Homo sapiens sapiens* could then be reserved for a future type of human that is able to meet defined criteria such as, for example, living within the planet's boundaries (*Nature* **461,** 472–475; 2009).

**Julian Cribb** *Nicholls, Australian Capital Territory, Australia. jcribb@work.netspeed.com.au*

## Rising ivory prices threaten elephants

Kenya's premier Samburu elephant population is the focus of a distressing surge in ivory poaching, coincident with an increase in illegal trading of ivory. This ivory is mainly destined for China (see go.nature.com/czac3x). Effective protection of elephants depends partly on more conservation investment, but mainly on stemming the demand for ivory and eliminating black-market trade — actions that mandate leadership from and cooperation with China.

The Samburu elephants are one of the world's best-studied populations. Intensive monitoring has revealed that more have been poached in the past 2.5 years than in the previous 11 years. The highest poaching rates ever were recorded in the first 5 months of this year.

Poaching of males has resulted in a population with more than double the number of females to males. Poachers are now also targeting adult females, resulting in the loss of one or more in most families and leaving roughly one in five groups with no mature females. The number of orphans in the population is increasing rapidly.

These changes correlate with a near tripling of the total number of seizures of illegal ivory in or coming from Kenya (see go.nature.com/k9bkwy) and with rising ivory prices. Local black-market ivory prices around Samburu have more than doubled since 2007, and are an order of magnitude greater than in 1990. At current local prices, the ivory of the largest male elephant poached in the Samburu population is equivalent to 1.5 years' salary for a wildlife ranger or 15 years'


© 2011 Macmillan Publishers Limited. All rights reserved

salary for an unskilled worker.

Ivory demand and prices have reached a point at which poachers are willing to target well-protected, closely monitored populations. With many poorly protected, soft-target elephant populations now over-harvested (P. Bouche *et al. PLoS ONE* **6,** e20619; 2011), the pressure on the Samburu elephant population may be a harbinger of what is to come for Africa's protected areas.

**George Wittemyer** *Colorado State University, Fort Collins, USA. g.wittemyer@colostate.edu*
**David Daballen, Iain Douglas-Hamilton** *Save the Elephants, Nairobi, Kenya.*

## Endangered cultures need protection too

Many strategies are used to protect wild plants and animals, but we are much more lax when it comes to ancestral cultures that are threatened or endangered worldwide.

We should recognize the few humans still living in isolation as infallible indicators of true wilderness. Both are vanishingly rare and our responsibilities towards them are tremendous. Any country that can boast the presence of an uncontacted group is truly fortunate and should realize that this represents an unusual opportunity to save two irreplaceable assets — a unique people and a wild place — for the price of one.

Ecuador is one such country. The Tagaeri and the Taromenane clans of the Waorani nation, voluntarily separated from Western civilization in space and time, roam the lowland rainforests of Yasuní, a region that has been documented as housing the greatest species richness on the planet.

In eastern Ecuador, as elsewhere, valuable resources located within indigenous peoples' homelands put them and their cultures at risk. Resorting to violence to push them aside has rightly come to be considered abominable. Financial means (from job offers and donations to promises of improved standards of living) are more acceptable and widely used to pacify native peoples on contact, or simply to get them out of the way — always under the pretext of a country's greater economic good.

A small isolated group of humans with no vote or voice is as valid for preservation as an endangered species. A country such as Ecuador has everything to gain by protecting both intact ecosystems and cultures through hands-off policies that are akin to those used for endangered species.

**Kelly Swing** *San Francisco de Quito University, Quito, Ecuador. kswing@usfq.edu.ec*

## US Congress right to halt GM salmon

You are incorrect in assuming that US federal-agency decisions on genetically modified (GM) organisms are always based on sound science and that Congress might be undermining such decisions (*Nature* **475,** 5; 2011).

In the case of GM salmon, the Advisory Committee of the US Food and Drug Administration (FDA) scolded the FDA and the salmon-production company AquaBounty of Waltham, Massachusetts, for overlooking important scientific issues during the approval process. Criticisms included inadequate sample sizes, incomplete data, questionable culling practices, troubling physical abnormalities, a failed test for allergenicity, and poor environmental and scientific assessments. So far, the agency has not corrected any of these oversights.

Furthermore, the FDA has never promulgated mandatory regulations or amended existing ones to cover GM animals. Instead, it announced in its non-binding 2009 *Guidance for Industry* (see go.nature.com/nvcvnx) that it would approve such animals under its existing 'new animal' drug law. As a result, the FDA — which has insufficient scientific expertise in assessing fisheries and environmental risk — is giving itself the task of reviewing and assessing GM fish. This is hardly a prescription for sound science.

Rather than blaming Congress, we should be glad that it has stepped in to address this serious economic, human-health and environmental issue.

**Andrew Kimbrell, Colin O'Neil** *Center for Food Safety, Washington DC, USA. kimbrell@icta.org*

## Balance the blend of food and fuel

Biofuels from non-food sources (lignocellulose, for example) are seen as a superior alternative to those from crops such as grain and sugar — fuels mistakenly assumed to have caused the 2008 spike in food prices. Competition for farmland continues to stoke the food-versus-biofuel debate, but a more inclusive and less simplistic framework is needed.

In a world where 33% of the population is overweight, 17% is undernourished and 40% of the food in developed countries is wasted, other aspects of food security should be considered. These include access to food (through governance, cost and distribution), nutritional adequacy and consumer behaviour.

Farmers already respond to policy and economic signals such as subsidies and renewable-energy mandates by diversifying their production systems. Integrating food crops, bioenergy feedstock and carbon farming could be an important strategy to counter economic and environmental variability.

Increasing global demand for food and energy means that discussions of land use should include the whole energy sector, not just the minor biofuel component. Other energy industries such as coal-bed methane producers, as well as urbanization, are also competing for arable land and potable water.

Expanding the debate in this way will help to integrate agricultural, energy and natural-resource policies to meet the needs of governments, producers and consumers in achieving food and fuel security, sustainable management of natural resources and reduced greenhouse-gas emissions.

**Luis C. Rodriguez, Deborah O'Connell** *CSIRO Ecosystem Sciences, Crace, Australian Capital Territory, Australia. deborah.o'connell@csiro.au*

## Nanomaterials need flexible regulation

Andrew Maynard rightly points out that 'one-size-fits-all' definitions are potentially harmful as they cannot capture the range of risks associated with a diversity of nanomaterials (*Nature* **475,** 31; 2011). But in arguing that regulators should instead work with a list of functional attributes to develop proper science-based regulations, he overlooks the fact that knowledge of such attributes, and internationally agreed methods for acquiring it, will take years of research to accumulate.

Given this situation, adopting a functional definition arguably means that, in theory, the most scientifically sound approach to regulation would be to impose a moratorium on the use of nanomaterials until the relevant trigger-point attributes can be established. There would still be the problem of defining the point in the life cycle of a nanoproduct at which such attributes should be measured, given that they would be likely to change with time and context.

The concept of adaptive regulation itself needs to be flexible enough to recognize the limitations of scientific understanding and its legitimacy as a basis for policy.

**Fern Wickson** *GenØk Centre for Biosafety, Tromsø, Norway. fern.wickson@genok.org*

**CONTRIBUTIONS**
Correspondence may be sent to **correspondence@nature.com** after consulting the author guidelines at **go.nature.com/cMCHno**.



© 2011 Macmillan Publishers Limited. All rights reserved

# John Marburger
## *(1941–2011)*

Physicist and longest-serving US presidential science adviser.

Physicist, university president and science adviser to former US president George W. Bush, John H. Marburger III often placed himself at the centre of difficult situations. He won widespread respect and admiration for his cool temperament, fairness and humanity. A colleague characterized him as "one of the most effective and farsighted scientific administrators of the postwar period". Marburger also faced intense criticism.

Marburger, who died on 28 July 2011 aged 70, was born on Staten Island, New York, and grew up in Severna Park, Maryland. In 1962, he completed a bachelor's degree in physics at Princeton University, New Jersey, going on to earn a PhD in applied physics at Stanford University in Palo Alto, California, in 1967. He joined the faculty of the University of Southern California in Los Angeles, where his research focused on non-linear and quantum optics, but he soon began to take on a series of challenging administrative posts.

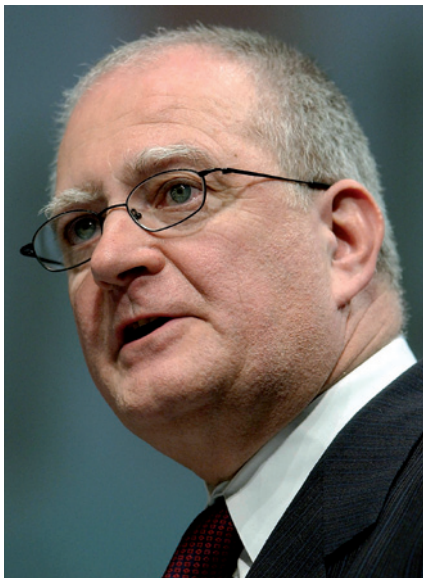
In 1972, he became chair of the physics department and four years later was appointed dean. In 1980, at the age of 39, he was made president of Stony Brook University in New York. Not long after, Mario Cuomo, the state governor, asked Marburger to chair a commission to give scientific advice on the future of the Shoreham nuclear power facility, then under construction but with an uncertain future due to technical problems, cost overruns and public opposition.

Marburger's experiences on the Shoreham commission shaped his thinking about the role of science in society and scientific advice in politics. Through interacting with the public, he quickly realized that decision-making involving nuclear power went beyond technical questions or whether facts were right or wrong. "There's a lot of sociology involved," he said. "If it governs the way people behave, then the factual becomes almost irrelevant."

Despite winning approval for his even-handedness, some criticized him for allowing too much public input into the process. Countering that "there were democratic considerations", Marburger argued that this commission was an opportunity for the public to feel that they were being heard.

After stepping down as president at Stony Brook, Marburger returned to research. Just four years later, in 1998, he became director of Brookhaven National Laboratory in Upton, New York, then engulfed by public criticism following a release of radioactive tritium. Marburger took on the job of rebuilding public trust in the institution through openness and transparency. He invited external regulators to sample the water and declared that no data would be withheld. "We can't stonewall," he said.

Marburger turned around Brookhaven's reputation, receiving accolades from environmental groups long critical of the lab. His commitment to public engagement was central to his management approach. When the UK newspaper *The Sunday Times* hyped the new Relativistic Heavy Ion Collider at Brookhaven with the headline "Big Bang machine could destroy Earth", Marburger formed an expert commission to allay public concern and foster public interest.



D. FAVRE/ARC/REUTERS

### THE BUSH YEARS

With his demonstrated flair for science, management and public engagement in highly politicized settings, his appointment as science adviser to Bush seemed a natural career step, and one that would be welcomed by the scientific community. Yet, from the first hints that Marburger was being considered for the position, he faced criticism from scientists for his defences of (or for being a part of) the Republican Bush administration.

Unlike previous science advisers but consistent with his commitment to transparency, Marburger declared his political affiliation as a Democrat. Under Bush, he was a consummate public servant, working behind the scenes from 2001 to 2009 as the longest-serving US science adviser. He briefed the president and agencies following the September 2001 terrorist attacks and anthrax scares, and helped to orient US research and development programmes in the nascent Department of Homeland Security. He also oversaw a massive increase in medical-research spending and the response to the loss of space shuttle *Columbia* in 2003.

During a public interview in 2005, I pressed Marburger to react to allegations that there was a systematic effort under way at the administration to undermine scientific integrity across the government. Marburger responded that such allegations — summarized at the time in reports by the Union of Concerned Scientists and Congressman Henry Waxman (Democrat, California) — amounted to little more than a conspiracy theory. "I don't give a lot of mental space to it," he said, "I don't think that it is seriously affecting science in the United States."

Although partisans continue to argue over the Bush administration's legacy, time has shown Marburger's judgement to be correct on at least two counts. First, US science remains strong. During his tenure, federal research budgets grew by the largest amount in a generation. Second, after his departure, issues of scientific integrity continue to be problematic under the administration of President Barack Obama, and remain beyond the scope of the science adviser's office.

In conversations I had with him over the past six years, as a subject of my research and later as a collaborator, Marburger never expressed resentment towards his critics. Earlier in his career he explained that "anger is not a useful reaction except symbolically, in a public-relations sense". How did he keep so calm? In his own words, he was an "incurable optimist" who found solace in science: "physics has been the main stabilizer of my life." In the last weeks of his life, Marburger finished a book on quantum physics, *Constructing Reality* which will be published by Cambridge University Press in September. ■

**Roger Pielke Jr** *is at the Center for Science and Technology Policy Research, University of Colorado, Boulder, Colorado 80309, USA. He first met Marburger while working on the book* Presidential Science Advisors: Perspective and Reflections on Science, Policy and Politics *(Springer, 2010).*
*e-mail: pielke@colorado.edu*



© 2011 Macmillan Publishers Limited. All rights reserved

FLUID MECHANICS

# When shape matters

**The well-known boundaries of coffee stains are caused by the outward flow of particles suspended in the liquid. Experiments show that ellipsoidal particles can prevent the formation of such boundaries.** SEE LETTER P.308

JAN VERMANT

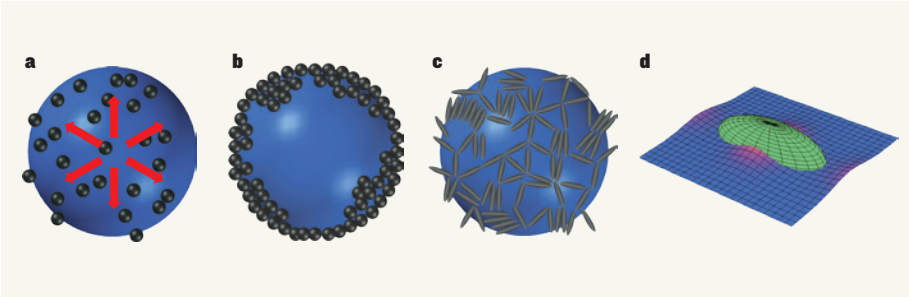



**Figure 1 | Preventing the coffee-ring effect.** **a**, When a liquid drop (blue; top view) containing suspended particles (black spheres) dries up, liquid flows from the interior of the bulk liquid towards the edge, carrying the particles with it (red arrows). **b**, Accumulation of the particles near the edge of the drop leads to the formation of a stain with a pronounced boundary (the coffee-ring effect). **c**, Yunker *et al.*[1] show that this can be prevented by having self-assembling ellipsoidal particles (grey) at the liquid–air interface. **d**, The local curvature of the ellipsoidal particle's surface (green) and the wetting of the particle by the fluid force the interface (blue and pink) to become deformed, pulling the liquid up near the middle of the ellipsoid and pushing it down near its tips.

Anyone who has spilt coffee is familiar with the characteristic ring of deposit that remains after the liquid in the circular stain has evaporated — a feature known as the coffee-ring effect. Whereas this may simply seem like one of nature's annoying tricks, the underlying physics causes problems for a number of technological applications, such as coating techniques or inkjet printing. On page 308 of this issue, Yunker *et al.*[1] demonstrate how the infamous coffee-ring effect can be easily prevented. When the shape of the particles suspended in the drop deviates sufficiently from a sphere, surface tension efficiently traps the particles at the liquid–air interface, where they subsequently assemble. The resulting loosely packed, interfacial particle assemblies provide a mechanical resistance at the interface that stabilizes the drop drying process.

The coffee-ring effect in a dried stain is observed for many types of suspended material. The mechanism underlying the accumulation of deposit near the rim of a drop was identified[2] a few years ago. The edge of the drop can become pinned to the substrate, for example because of the substrate's surface roughness. This causes the drop radius to stay constant as it evaporates. Because liquid lost by evaporation around the edge must be replenished, surface tension causes the liquid to flow from the interior towards the edge (Fig. 1a). Material suspended in the liquid is therefore carried along towards the rim of the drop, accumulating there and forming the characteristic ring-like stain (Fig. 1b).

To counteract surface tension and thus prevent the coffee-ring effect, Yunker *et al.*[1] exploited particle-shape effects to build a particle network at the interface between the air and the water. This leads to an increase in the interfacial viscosity and prevents the suspended particles from reaching the drop edge. It was Benjamin Franklin who first studied how increasing the interfacial viscosity can stabilize flows within the bulk of a liquid, by spreading oil on water and stilling the waves and ripples on a London pond[3].

Yunker and colleagues have not used an additional substance such as oil, but instead demonstrate how the increase in interfacial flow resistance can be built into a suspension. It is here that shape becomes relevant, because, for ellipsoidal particles, surface tension can be turned against itself. First, it causes the particles to become strongly attached to the air–water interface as they are carried to the drop edge by the outward flow in the thinning drop. Second, surface tension is responsible for the lateral capillary forces that push the particles together once they are at the interface, causing the ellipsoids to assemble and form open, aggregated structures (Fig. 1c).

The lateral capillary forces are linked to yet another breakfast-related phenomenon, known as the Cheerios effect[4], which refers to the tendency of small floating objects to attract one another, as seen for example in the breakfast cereal after which this effect is named. In the case of the cereal floating on the surface of milk, the weight of the particles causes them to deform the air–milk interface. To minimize the overall distortion of the interface, surface tension forces the particles together.

However, for sufficiently small particles, gravity can no longer deform the interface, and yet shape effects can be cleverly exploited, as Yunker and co-workers demonstrate. When particles with an ellipsoidal shape are partially immersed in water, because of the local curvature of the particles' surface, the wetting of the particle by water deforms the air–water interface, pulling the liquid up near the middle of the ellipsoid and pushing it down near its tips (Fig. 1d). The ensuing local change in the shape of the liquid interface leads to lateral capillary forces, which have been demonstrated to be both long-range and strong[5]. Particle roughness or chemical heterogeneity, as opposed to ellipsoidal particle shape, might be used to obtain similar lateral capillary forces. But it is not known what amount of roughness leads to comparable interfacial viscosity and uniform particle deposition.

In their experiments, Yunker and colleagues compared the behaviour of micrometre-sized polystyrene latex spheres with that of the same particles stretched into ellipsoids. They produced drops of aqueous suspensions at concentrations of 0.01–20 per cent by volume and studied them using video microscopy. They found that, if the aspect ratio of the large axis to the minor axis of the ellipsoid is larger than 2.5, the suspensions deposit as uniform stains. By using confocal microscopy, they then obtained direct evidence that, for ellipsoids, particle assembly occurs at the interface.

The effects of the addition of small amounts of surfactant to the suspensions provided further support for the authors' hypothesis that interfacial assembly of ellipsoids is responsible for stopping the ring effect. Surfactants reduce surface tension and hence negatively affect the efficiency of particle self-assembly, restoring


© 2011 Macmillan Publishers Limited. All rights reserved

the coffee-ring effect. However, the effect of surfactants may be more complicated than shown for the particular concentration used, given their own role in potentially suppressing the coffee-ring effect through the manipulation of gradients in surface tension[6].

Interestingly, the authors find that adding even small amounts (as little as 0.015%) of ellipsoids to a suspension of spheres is enough to suppress the formation of the ring. This finding should motivate research into the interfacial viscosity and elasticity of these mixed systems.

Yunker and colleagues' results may be of great practical importance. First, building an interfacial resistance to bulk flow by exploiting the effects of particle shape can be a generic route to stabilizing a number of processes in which the coffee-ring effect is problematic. For example, minor amounts of non-spherical particles could provide the means to achieve uniform coating of thin films, which would not otherwise be possible. Second, it is worth considering how wetting and de-wetting flows will be affected when a high interfacial flow resistance is present — this occurs when a liquid is in direct contact with a solid, and so is relevant to paints, coatings and lubricants, for example. The use of ellipsoidal particles provides an attractive way to investigate these phenomena.

The study's results[1] may even be relevant to things as mundane as the food we eat or the consumer products we use. Many of those come in the form of foams or emulsions — think of creams and lotions used as cosmetics or pharmaceuticals, hair-styling products and shaving foam. The flow in thin films, which causes the rupture of bubbles in foam or drops in emulsions, resembles to some extent that of the drop drying process. A strong interfacial resistance to flow can be used to stabilize emulsions, and ellipsoidal particles have been shown[7] to be extremely efficient at doing this. Exploiting particle shape may thus be a generic way to 'still' certain types of bulk flows driven by surface tension and to further our understanding of interfacial fluid mechanics. ■

**Jan Vermant** *is in the Department of Chemical Engineering, Katholieke Universiteit Leuven, 3001 Leuven, Belgium.*
*e-mail: jan.vermant@cit.kuleuven.be*

1. Yunker, P. J., Still, T., Lohr, M. A. & Yodh, A. G. *Nature* **476,** 308–311 (2011).
2. Deegan, R. D. *et al. Nature* **389,** 827–829 (1997).
3. Franklin, B. *et al. Phil. Trans.* **64,** 445–460 (1774).
4. Vella, D. & Mahadevan, L. *Am. J. Phys.* **73,** 817–825 (2005).
5. Loudet, J. C., Alsayed, A. M., Zhang, J. & Yohd, A. G. *Phys. Rev. Lett.* **94,** 018301 (2005).
6. Hu, H. & Larson, R. G. *J. Phys. Chem. B* **110,** 7090–7094 (2006).
7. Madivala, B., Vandebril, S., Fransaer, J. & Vermant, J. *Soft Matter* **5,** 1717–1727 (2009).

STEM CELLS

# Orphan receptors find a home

**The protein Lgr5 has been valuable as the undisputed marker of intestinal and other stem cells. It emerges that Lgr5 and its relatives also have essential signalling roles of relevance to health and disease.** SEE ARTICLE P.293

WALTER BIRCHMEIER

If you want to find a human stem cell in a tissue, good luck: they are usually rare and are mixed with other specialized cell types. Yet identifying stem cells is an essential step in isolating them for research or therapeutic purposes. This is usually done by searching for markers that occur on or within only a specific stem-cell type. The identification of such markers can also have a secondary payoff, because molecules exclusive to a stem cell probably have some vital function. But, until now, the functions of Lgr5 — a cell-surface protein that marks certain adult stem cells — and other Lgr proteins have remained unclear. Two papers[1,2], including one by de Lau *et al.* on page 293 of this issue, report that the R-spondin growth factors bind to Lgr proteins and mediate activation of the Wnt signalling pathway, which has roles in animal development, as well as in diseases such as cancer[3].

Lgr proteins are cell-surface receptors. Lgr1, Lgr2 and Lgr3 are hormone receptors that signal through G proteins. But Lgr4, Lgr5 and Lgr6 have been considered 'orphan' receptors because, despite extensive searching, their activators remained mysterious. Nonetheless, Lgr4 is known to have an essential role in the development of the male reproductive tract, as well as the eyelids, hair and bone. Mice lacking Lgr5 die shortly after birth of malformation of the tongue and lower jaw[4]. And Lgr6, like Lgr5, is a specific stem-cell marker in several epithelial layers: whereas Lgr5 marks proliferative stem cells in the gastrointestinal tract, Lgr6 specifically occurs on the multi-potential stem cells of the skin[5,6].

Writing in *Proceedings of the National Academy of Sciences*, Carmon *et al.*[1] draw mostly on classical biochemical approaches to show that, in cultured cells, R-spondins (which are secreted agonists of the Wnt pathway) can bind directly to Lgr4 and Lgr5. De Lau *et al.*[2] show that R-spondins bind not only to Lgr4 and Lgr5, but also to Lgr6, and further explore the outcome of R-spondin–Lgr interaction. They report that Lgr4 and Lgr5 specifically associate with Frizzled and LRPs — proteins that are activated by the extracellular Wnt molecules and that then trigger canonical Wnt signalling to increase gene expression (Fig. 1).

R-spondins are produced by the cells of deep, well-like structures called intestinal crypts, and strongly promote cell division in Lgr5-positive stem cells[7], which also reside in the crypts. Intestinal stem cells are responsible for the formation of new gut tissue, and so generate substantial numbers of intestinal epithelial cells — amounting to several hundred grams per day. Therefore, any disruption of stem-cell maintenance, growth or differentiation is expected to lead to the development of severe intestinal defects and/or cancer. De Lau *et al.* report that, on deletion of the genes encoding Lgr4 and Lgr5, mice lose their intestinal stem cells.

Wnt signals help regulate cell division and specialization throughout the body during

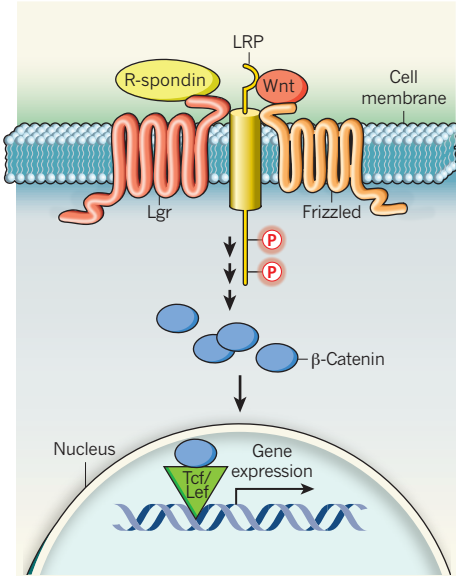



**Figure 1 | Cooperative signalling within stem cells.** R-spondins activate Lgr membrane proteins, which are also specific stem-cell markers. Lgr proteins then recruit the LRP–Frizzled receptor complex, which binds to Wnt ligands, reinforcing Wnt signalling following phosphorylation (P) of LRP. A series of steps that follow stabilizes β-catenin, which is translocated to the nucleus and, together with the Tcf/Lef family of transcription factors, induces gene expression.



© 2011 Macmillan Publishers Limited. All rights reserved

the coffee-ring effect. However, the effect of surfactants may be more complicated than shown for the particular concentration used, given their own role in potentially suppressing the coffee-ring effect through the manipulation of gradients in surface tension[6].

Interestingly, the authors find that adding even small amounts (as little as 0.015%) of ellipsoids to a suspension of spheres is enough to suppress the formation of the ring. This finding should motivate research into the interfacial viscosity and elasticity of these mixed systems.

Yunker and colleagues' results may be of great practical importance. First, building an interfacial resistance to bulk flow by exploiting the effects of particle shape can be a generic route to stabilizing a number of processes in which the coffee-ring effect is problematic. For example, minor amounts of non-spherical particles could provide the means to achieve uniform coating of thin films, which would not otherwise be possible. Second, it is worth considering how wetting and de-wetting flows will be affected when a high interfacial flow resistance is present — this occurs when a liquid is in direct contact with a solid, and so is relevant to paints, coatings and lubricants, for example. The use of ellipsoidal particles provides an attractive way to investigate these phenomena.

The study's results[1] may even be relevant to things as mundane as the food we eat or the consumer products we use. Many of those come in the form of foams or emulsions — think of creams and lotions used as cosmetics or pharmaceuticals, hair-styling products and shaving foam. The flow in thin films, which causes the rupture of bubbles in foam or drops in emulsions, resembles to some extent that of the drop drying process. A strong interfacial resistance to flow can be used to stabilize emulsions, and ellipsoidal particles have been shown[7] to be extremely efficient at doing this. Exploiting particle shape may thus be a generic way to 'still' certain types of bulk flows driven by surface tension and to further our understanding of interfacial fluid mechanics. ■

**Jan Vermant** *is in the Department of Chemical Engineering, Katholieke Universiteit Leuven, 3001 Leuven, Belgium.*
*e-mail: jan.vermant@cit.kuleuven.be*

1. Yunker, P. J., Still, T., Lohr, M. A. & Yodh, A. G. *Nature* **476,** 308–311 (2011).
2. Deegan, R. D. *et al. Nature* **389,** 827–829 (1997).
3. Franklin, B. *et al. Phil. Trans.* **64,** 445–460 (1774).
4. Vella, D. & Mahadevan, L. *Am. J. Phys.* **73,** 817–825 (2005).
5. Loudet, J. C., Alsayed, A. M., Zhang, J. & Yohd, A. G. *Phys. Rev. Lett.* **94,** 018301 (2005).
6. Hu, H. & Larson, R. G. *J. Phys. Chem. B* **110,** 7090–7094 (2006).
7. Madivala, B., Vandebril, S., Fransaer, J. & Vermant, J. *Soft Matter* **5,** 1717–1727 (2009).

STEM CELLS

# Orphan receptors find a home

**The protein Lgr5 has been valuable as the undisputed marker of intestinal and other stem cells. It emerges that Lgr5 and its relatives also have essential signalling roles of relevance to health and disease.** SEE ARTICLE P.293

WALTER BIRCHMEIER

If you want to find a human stem cell in a tissue, good luck: they are usually rare and are mixed with other specialized cell types. Yet identifying stem cells is an essential step in isolating them for research or therapeutic purposes. This is usually done by searching for markers that occur on or within only a specific stem-cell type. The identification of such markers can also have a secondary payoff, because molecules exclusive to a stem cell probably have some vital function. But, until now, the functions of Lgr5 — a cell-surface protein that marks certain adult stem cells — and other Lgr proteins have remained unclear. Two papers[1,2], including one by de Lau *et al.* on page 293 of this issue, report that the R-spondin growth factors bind to Lgr proteins and mediate activation of the Wnt signalling pathway, which has roles in animal development, as well as in diseases such as cancer[3].

Lgr proteins are cell-surface receptors. Lgr1, Lgr2 and Lgr3 are hormone receptors that signal through G proteins. But Lgr4, Lgr5 and Lgr6 have been considered 'orphan' receptors because, despite extensive searching, their activators remained mysterious. Nonetheless, Lgr4 is known to have an essential role in the development of the male reproductive tract, as well as the eyelids, hair and bone. Mice lacking Lgr5 die shortly after birth of malformation of the tongue and lower jaw[4]. And Lgr6, like Lgr5, is a specific stem-cell marker in several epithelial layers: whereas Lgr5 marks proliferative stem cells in the gastrointestinal tract, Lgr6 specifically occurs on the multi-potential stem cells of the skin[5,6].

Writing in *Proceedings of the National Academy of Sciences*, Carmon *et al.*[1] draw mostly on classical biochemical approaches to show that, in cultured cells, R-spondins (which are secreted agonists of the Wnt pathway) can bind directly to Lgr4 and Lgr5. De Lau *et al.*[2] show that R-spondins bind not only to Lgr4 and Lgr5, but also to Lgr6, and further explore the outcome of R-spondin–Lgr interaction. They report that Lgr4 and Lgr5 specifically associate with Frizzled and LRPs — proteins that are activated by the extracellular Wnt molecules and that then trigger canonical Wnt signalling to increase gene expression (Fig. 1).

R-spondins are produced by the cells of deep, well-like structures called intestinal crypts, and strongly promote cell division in Lgr5-positive stem cells[7], which also reside in the crypts. Intestinal stem cells are responsible for the formation of new gut tissue, and so generate substantial numbers of intestinal epithelial cells — amounting to several hundred grams per day. Therefore, any disruption of stem-cell maintenance, growth or differentiation is expected to lead to the development of severe intestinal defects and/or cancer. De Lau *et al.* report that, on deletion of the genes encoding Lgr4 and Lgr5, mice lose their intestinal stem cells.

Wnt signals help regulate cell division and specialization throughout the body during

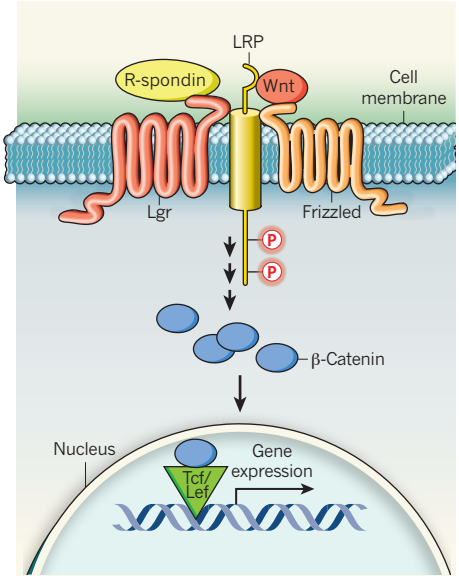



**Figure 1 | Cooperative signalling within stem cells.** R-spondins activate Lgr membrane proteins, which are also specific stem-cell markers. Lgr proteins then recruit the LRP–Frizzled receptor complex, which binds to Wnt ligands, reinforcing Wnt signalling following phosphorylation (P) of LRP. A series of steps that follow stabilizes β-catenin, which is translocated to the nucleus and, together with the Tcf/Lef family of transcription factors, induces gene expression.



© 2011 Macmillan Publishers Limited. All rights reserved

animal development and tissue regeneration. It is not surprising, therefore, that defects in the Wnt pathway have been linked to cancer progression in many organs, as well as to bone disease, limb, kidney and cardiac defects, diabetes and neurodegenerative disorders[3]. Moreover, mutations in R-spondins, in Lgr proteins and in components of canonical Wnt signalling in humans and mice produce similar defects in many tissues, including the intestine, kidneys, hair, bones, the vascular system, sexual organs and the spinal cord[3,4,6,8–11].

The two papers[1,2] imply that R-spondin signals augment Wnt signals to the intestinal stem cells. But to corroborate this will require an intimate knowledge of normal and abnormal Wnt signalling in various contexts. Nineteen secreted members of the Wnt family can bind to Frizzled receptors and to their associated LRP5 and LRP6 co-receptors at the membrane (Fig. 1). Successful activation leads to phosphorylation of LRP receptors, which eventually interferes with degradation of the crucial signalling molecule β-catenin. Consequently, β-catenin accumulates and translocates to the nucleus to alter — together with the Tcf/Lef family of transcription factors — the expression of a wide range of target genes.

Among the genes affected is c-*myc* — a gene commonly mutated in cancer. Activated Wnt/β-catenin signalling is also perceived to be essential for the cancer-associated processes of epithelial–mesenchymal transition and metastasis. Tumour cells that undergo epithelial–mesenchymal transition have similar characteristics to embryonic stem cells[12], which supports the idea that Wnt signalling mediates their 'stemness'.

Demonstration of a connection between R-spondins, Lgr proteins and Wnt signalling raises a question about its logistics. For R-spondin to convey its signal, it must bind to Lgr proteins, whereas Wnt agonists bind to the Frizzled–LRP complex. How do all these proteins come together to form a complex in the cell membrane (Fig. 1)? A detailed picture of these associations might not only show how they can be blocked or enhanced for therapeutic purposes, but also help to explain how stem cells fine-tune the signals that they receive.

But obtaining that picture will probably be challenging, because components of protein complexes seem to interact dynamically and promiscuously. Ultimately, an understanding of these processes is likely to require a detailed structural view of the proteins that are involved and the interactions between them. A first step would be to introduce targeted mutations into Lgr proteins — for example, to determine which specific regions of the molecule influence complex formation. ■

**Walter Birchmeier** *is at the Max Delbrück Center for Molecular Medicine, 13125 Berlin-Buch, Germany.*
*e-mail: wbirch@mdc-berlin.de*

1. Carmon, K. S., Gong, X., Lin, Q., Thomas, A. & Liu, Q. *Proc. Natl Acad. Sci. USA* **108,** 11452–11457 (2011).
2. de Lau, W. *et al. Nature* **476,** 293–297 (2011).
3. Klaus, A. & Birchmeier, W. *Nature Rev. Cancer* **8,** 387–398 (2008).
4. Morita, H. *et al. Mol. Cell. Biol.* **24,** 9736–9743 (2004).
5. Snippert, H. J. *et al. Science* **327,** 1385–1389 (2010).
6. Barker, N. *et al. Nature* **449,** 1003–1007 (2007).
7. Sato, T. *et al. Nature* **459,** 262–265 (2009).
8. Blaydon, D. C. *et al. Nature Genet.* **38,** 1245–1247 (2006).
9. Binnerts, M. E. *et al. Proc. Natl Acad. Sci. USA* **104,** 14700–14705 (2007).
10. Parma, P. *et al. Nature Genet.* **38,** 1304–1309 (2006).
11. Kazanskaya, O. *et al. Development* **135,** 3655–3664 (2008).
12. Ben-Porath, I. *et al. Nature Genet.* **40,** 499–507 (2008).

ASTRONOMY

# Unlocking the secrets of the giant blobs

**Polarized emission has been detected from the largest Lyman-α gas cloud, known as blob 1. This result strongly suggests that such clouds are powered by a central source of ionizing radiation.** SEE LETTER P.304

RICHARD BOWER

Sometimes the most fruitful discoveries are the most unexpected. The discovery of Lyman-α blobs, puzzling giant gas clouds seen in the young Universe, definitely falls into this class. If astronomers could only agree on their power source, these objects would offer a tantalizing insight into the formation of young galaxies. But their nature is controversial. On page 304 of this issue, Hayes *et al.*[1] report the first detection of polarized emission from the largest known cloud, Lyman-α blob 1. Their finding strongly suggests that the cloud is visible because of the scattering of photons in an extended gaseous halo.

The story begins in 2000, when Steidel and collaborators[2] undertook a survey to identify faint galaxies displaying narrow spectral emission lines. Their study was intended to count the abundance of galaxies in the young Universe, but they were blown away by an unexpected discovery. Their survey revealed two huge, glowing gas clouds. These are seen 11 billion years in the past (when the Universe was one-sixth of its current age), at an epoch when most of the stars in the Universe had already formed. The clouds are immense, a hundred times larger than the galaxies they contain. Steidel *et al.*[2] were lost for words

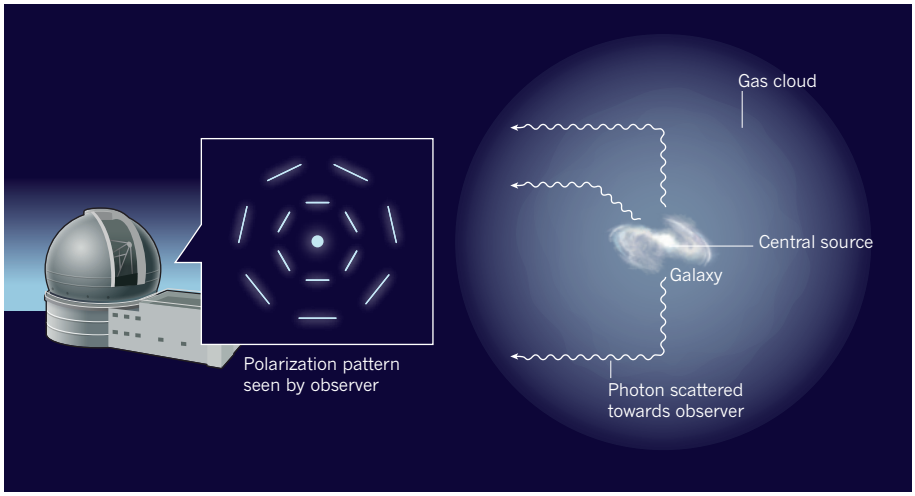



**Figure 1 | Polarized emission from Lyman-α blobs.** Lyman-α photons are emitted from a young galaxy at the centre of a gas cloud. Photons usually escape only after they have undergone Doppler scattering. On average, photons seen by an observer on Earth have been scattered at right angles to their direction of emission, inducing a small but just measurable polarization. The observer will see rings of polarization centred on the original source. By contrast, if the photons originate from many randomly located sources, no polarization pattern will be evident. Hayes *et al.*[1] have detected rings of polarized emission from the largest Lyman-α blob known.


© 2011 Macmillan Publishers Limited. All rights reserved

animal development and tissue regeneration. It is not surprising, therefore, that defects in the Wnt pathway have been linked to cancer progression in many organs, as well as to bone disease, limb, kidney and cardiac defects, diabetes and neurodegenerative disorders[3]. Moreover, mutations in R-spondins, in Lgr proteins and in components of canonical Wnt signalling in humans and mice produce similar defects in many tissues, including the intestine, kidneys, hair, bones, the vascular system, sexual organs and the spinal cord[3,4,6,8–11].

The two papers[1,2] imply that R-spondin signals augment Wnt signals to the intestinal stem cells. But to corroborate this will require an intimate knowledge of normal and abnormal Wnt signalling in various contexts. Nineteen secreted members of the Wnt family can bind to Frizzled receptors and to their associated LRP5 and LRP6 co-receptors at the membrane (Fig. 1). Successful activation leads to phosphorylation of LRP receptors, which eventually interferes with degradation of the crucial signalling molecule β-catenin. Consequently, β-catenin accumulates and translocates to the nucleus to alter — together with the Tcf/Lef family of transcription factors — the expression of a wide range of target genes.

Among the genes affected is c-*myc* — a gene commonly mutated in cancer. Activated Wnt/β-catenin signalling is also perceived to be essential for the cancer-associated processes of epithelial–mesenchymal transition and metastasis. Tumour cells that undergo epithelial–mesenchymal transition have similar characteristics to embryonic stem cells[12], which supports the idea that Wnt signalling mediates their 'stemness'.

Demonstration of a connection between R-spondins, Lgr proteins and Wnt signalling raises a question about its logistics. For R-spondin to convey its signal, it must bind to Lgr proteins, whereas Wnt agonists bind to the Frizzled–LRP complex. How do all these proteins come together to form a complex in the cell membrane (Fig. 1)? A detailed picture of these associations might not only show how they can be blocked or enhanced for therapeutic purposes, but also help to explain how stem cells fine-tune the signals that they receive.

But obtaining that picture will probably be challenging, because components of protein complexes seem to interact dynamically and promiscuously. Ultimately, an understanding of these processes is likely to require a detailed structural view of the proteins that are involved and the interactions between them. A first step would be to introduce targeted mutations into Lgr proteins — for example, to determine which specific regions of the molecule influence complex formation. ■

**Walter Birchmeier** *is at the Max Delbrück Center for Molecular Medicine, 13125 Berlin-Buch, Germany.*
*e-mail: wbirch@mdc-berlin.de*

1. Carmon, K. S., Gong, X., Lin, Q., Thomas, A. & Liu, Q. *Proc. Natl Acad. Sci. USA* **108,** 11452–11457 (2011).
2. de Lau, W. *et al. Nature* **476,** 293–297 (2011).
3. Klaus, A. & Birchmeier, W. *Nature Rev. Cancer* **8,** 387–398 (2008).
4. Morita, H. *et al. Mol. Cell. Biol.* **24,** 9736–9743 (2004).
5. Snippert, H. J. *et al. Science* **327,** 1385–1389 (2010).
6. Barker, N. *et al. Nature* **449,** 1003–1007 (2007).
7. Sato, T. *et al. Nature* **459,** 262–265 (2009).
8. Blaydon, D. C. *et al. Nature Genet.* **38,** 1245–1247 (2006).
9. Binnerts, M. E. *et al. Proc. Natl Acad. Sci. USA* **104,** 14700–14705 (2007).
10. Parma, P. *et al. Nature Genet.* **38,** 1304–1309 (2006).
11. Kazanskaya, O. *et al. Development* **135,** 3655–3664 (2008).
12. Ben-Porath, I. *et al. Nature Genet.* **40,** 499–507 (2008).

ASTRONOMY

# Unlocking the secrets of the giant blobs

**Polarized emission has been detected from the largest Lyman-α gas cloud, known as blob 1. This result strongly suggests that such clouds are powered by a central source of ionizing radiation.** SEE LETTER P.304

RICHARD BOWER

Sometimes the most fruitful discoveries are the most unexpected. The discovery of Lyman-α blobs, puzzling giant gas clouds seen in the young Universe, definitely falls into this class. If astronomers could only agree on their power source, these objects would offer a tantalizing insight into the formation of young galaxies. But their nature is controversial. On page 304 of this issue, Hayes *et al.*[1] report the first detection of polarized emission from the largest known cloud, Lyman-α blob 1. Their finding strongly suggests that the cloud is visible because of the scattering of photons in an extended gaseous halo.

The story begins in 2000, when Steidel and collaborators[2] undertook a survey to identify faint galaxies displaying narrow spectral emission lines. Their study was intended to count the abundance of galaxies in the young Universe, but they were blown away by an unexpected discovery. Their survey revealed two huge, glowing gas clouds. These are seen 11 billion years in the past (when the Universe was one-sixth of its current age), at an epoch when most of the stars in the Universe had already formed. The clouds are immense, a hundred times larger than the galaxies they contain. Steidel *et al.*[2] were lost for words

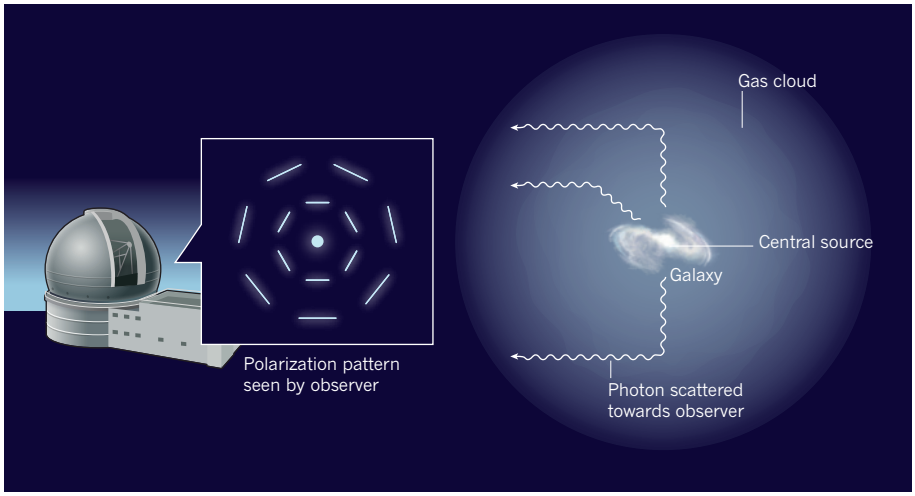



**Figure 1 | Polarized emission from Lyman-α blobs.** Lyman-α photons are emitted from a young galaxy at the centre of a gas cloud. Photons usually escape only after they have undergone Doppler scattering. On average, photons seen by an observer on Earth have been scattered at right angles to their direction of emission, inducing a small but just measurable polarization. The observer will see rings of polarization centred on the original source. By contrast, if the photons originate from many randomly located sources, no polarization pattern will be evident. Hayes *et al.*[1] have detected rings of polarized emission from the largest Lyman-α blob known.


© 2011 Macmillan Publishers Limited. All rights reserved

PSYCHOLOGY

# When it's fair to share

Humans have at least two characteristics to be proud of. Compared with other great apes, we are not only better collaborators but also more generous. On page 328 of this issue, Hamann *et al.* explore the connection between these two traits (K. Hamann *et al. Nature* **476,** 328–331; 2011).

The age at which humans first show tendencies towards sharing has been investigated extensively. Three-to-five-year-olds seem to divide any windfall resources unfairly, keeping something they have obtained without effort mostly to themselves. After that, however, they become more inclined to share windfalls equally.

Hamann and colleagues wondered whether children would behave differently if the resource had been obtained through joint effort. They studied pairs of 2- and 3-year-old children when, in each case, one child had greater access to the resources (to three of four toys) and so could choose whether to share them with their partner. The toys could be obtained in three ways: through the collaborative efforts of both children; freely, with no effort by either; or through independent work.

The authors found that 3-year-olds, but not 2-year-olds, in the position of power opted to share more often (approximately 75% of the time) after a collaborative effort than after parallel work (about 25%). So, although the toddlers may not have acquired a sense of 'reward in return for work', they understood that resource distribution should depend on how the resource was obtained.

Is this link between collaboration and fairness uniquely human? If other primates don't appreciate the value of collaboration as much as we do, they would be expected to be indifferent to the idea of allocating resources depending on how they were acquired.



DESIGN PICS INC/REX FEATURES

Testing this hypothesis, Hamann *et al.* found that, when in a position of advantage, chimpanzees sometimes shared, but this was independent of whether a food resource was obtained through collaboration or luck.

An explanation for this behaviour in chimps may be that, unlike in humans, foraging does not depend on collaboration. In other words, humans appreciate the importance of collaboration for survival and so invest in being included in future collaborative efforts by sharing existing resources. Luckily, we seem to learn this quite early on. **Sadaf Shadan**

— and these enormous clouds have become known simply as Lyman-α blobs.

Although Steidel and colleagues' initial measurements[2] detected only two bright emission clouds, follow-up work has shown that these are just the tip of the iceberg, with smaller clouds being seen around fainter galaxies. Of course, the key question is, what powers the blobs? If this could be answered for the largest systems, it could provide crucial insight into the formation of galaxies.

Ten years on, the puzzle of the blobs' identity remains. There are two front runners: that they are cooling radiation, released when filamentary gas falls rapidly onto early galaxies; or that the emission is light from a central burst of star formation or a black hole (an active galactic nucleus, AGN) that has been scattered by an extended halo of gas surrounding the system.

The first explanation has been gaining popularity for a while. Theoretical models suggest that galaxies form at the intersection of thin filaments of the 'cosmic web' of dark matter and gas that exist in the Universe. Gas flowing into the galaxies along these filaments must lose energy, providing an ample, distributed source of power[3,4]. Could the Lyman-α blobs be the observational confirmation of this prediction? By contrast, evidence for the second hypothesis is circumstantial. It comes from the frequently observed association[5,6] of bright Lyman-α blobs with AGNs and starburst galaxies, which form stars at exceptionally high rates. However, Hayes and colleagues' discovery[1] of polarized emission from the largest and brightest of the blobs now argues strongly in favour of the second option.

To appreciate the significance of this result, it is essential to understand what makes the emission from the blobs special. The emission is seen only at the wavelength corresponding to the Lyman-α electronic transition of hydrogen. The Lyman-α spectral line is created by a transition from the first excited state of hydrogen to the ground state. Because atoms in the ground state are common, the transition is resonant, meaning that a Lyman-α photon that is emitted by one atom is quickly absorbed by a neighbouring atom, promoting it to the excited state. This atom quickly decays to the ground state, releasing another Lyman-α photon, which will be absorbed by another atom, and so on.

It is therefore difficult for a Lyman-α photon to escape unless the medium is very highly ionized or the random motions of the atoms cause the photon to undergo Doppler shifting to a higher or lower frequency, so that the frequency of the photon no longer matches the Lyman-α transition of a hydrogen atom. Although this scattering in frequency leads to the photon's escape, it also leaves an imprint on the polarization of the photon because escaping photons are, on average, scattered perpendicular to their initial direction of travel[7]. A similar process is responsible for polarizing the light reflected from the surface of water, and is why many sunglasses contain polarized lenses. Reducing this complex problem to the simplest level, rings of polarized emission would be expected to be observed if the photons originate from a central source, whereas unpolarized emission would be expected if the source of the photons is widely distributed.

In Hayes and colleagues' technically impressive study[1] of the emission from the brightest of the Lyman-α blobs, the polarization pattern consists of concentric rings around the likely centre of the emission (Fig. 1), in good agreement with the predictions for scattering from a central source. If this result can be generalized to all Lyman-α blobs, it opens up the intriguing possibility of using the properties of the scattered radiation to probe the gas haloes of young galaxies. When combined with absorption measurements from background galaxies and AGNs, observations of extended Lyman-α blobs suggest a promising route to studying the previously invisible extended circum-galactic medium of galaxies[8]. This possibility is exciting, because stars and cold gas account for only a small fraction of the ordinary matter in the Universe: theoretical models predict that most ordinary matter will reside in an extended and hitherto undetectable reservoir. ■

**Richard Bower** *is at the Institute for Computational Cosmology, Department of Physics, University of Durham, Durham DH1 3LE, UK.*
*e-mail: r.g.bower@durham.ac.uk*

1. Hayes, M., Scarlata, C. & Siana, B. *Nature* **476,** 304–307 (2011).
2. Steidel, C. C. *et al. Astrophys. J.* **532,** 170–182 (2000).
3. Fardal, M. A. *et al. Astrophys. J.* **562,** 605–617 (2001).
4. Goerdt, T. *et al. Mon. Not. R. Astron. Soc.* **407,** 613–631 (2010).
5. Matsuda, Y. *et al. Astron. J.* **128,** 569–584 (2004).
6. Geach, J. E. *et al. Astrophys. J.* **700,** 1–9 (2009).
7. Dijkstra, M. & Loeb, A. *Mon. Not. R. Astron. Soc.* **386,** 492–504 (2008).
8. Steidel, C. C. *et al. Astrophys. J.* **736,** 160 (2011).



© 2011 Macmillan Publishers Limited. All rights reserved

PSYCHOLOGY

# When it's fair to share

Humans have at least two characteristics to be proud of. Compared with other great apes, we are not only better collaborators but also more generous. On page 328 of this issue, Hamann *et al.* explore the connection between these two traits (K. Hamann *et al. Nature* **476,** 328–331; 2011).

The age at which humans first show tendencies towards sharing has been investigated extensively. Three-to-five-year-olds seem to divide any windfall resources unfairly, keeping something they have obtained without effort mostly to themselves. After that, however, they become more inclined to share windfalls equally.

Hamann and colleagues wondered whether children would behave differently if the resource had been obtained through joint effort. They studied pairs of 2- and 3-year-old children when, in each case, one child had greater access to the resources (to three of four toys) and so could choose whether to share them with their partner. The toys could be obtained in three ways: through the collaborative efforts of both children; freely, with no effort by either; or through independent work.

The authors found that 3-year-olds, but not 2-year-olds, in the position of power opted to share more often (approximately 75% of the time) after a collaborative effort than after parallel work (about 25%). So, although the toddlers may not have acquired a sense of 'reward in return for work', they understood that resource distribution should depend on how the resource was obtained.

Is this link between collaboration and fairness uniquely human? If other primates don't appreciate the value of collaboration as much as we do, they would be expected to be indifferent to the idea of allocating resources depending on how they were acquired.



DESIGN PICS INC/REX FEATURES

Testing this hypothesis, Hamann *et al.* found that, when in a position of advantage, chimpanzees sometimes shared, but this was independent of whether a food resource was obtained through collaboration or luck.

An explanation for this behaviour in chimps may be that, unlike in humans, foraging does not depend on collaboration. In other words, humans appreciate the importance of collaboration for survival and so invest in being included in future collaborative efforts by sharing existing resources. Luckily, we seem to learn this quite early on. **Sadaf Shadan**

— and these enormous clouds have become known simply as Lyman-α blobs.

Although Steidel and colleagues' initial measurements[2] detected only two bright emission clouds, follow-up work has shown that these are just the tip of the iceberg, with smaller clouds being seen around fainter galaxies. Of course, the key question is, what powers the blobs? If this could be answered for the largest systems, it could provide crucial insight into the formation of galaxies.

Ten years on, the puzzle of the blobs' identity remains. There are two front runners: that they are cooling radiation, released when filamentary gas falls rapidly onto early galaxies; or that the emission is light from a central burst of star formation or a black hole (an active galactic nucleus, AGN) that has been scattered by an extended halo of gas surrounding the system.

The first explanation has been gaining popularity for a while. Theoretical models suggest that galaxies form at the intersection of thin filaments of the 'cosmic web' of dark matter and gas that exist in the Universe. Gas flowing into the galaxies along these filaments must lose energy, providing an ample, distributed source of power[3,4]. Could the Lyman-α blobs be the observational confirmation of this prediction? By contrast, evidence for the second hypothesis is circumstantial. It comes from the frequently observed association[5,6] of bright Lyman-α blobs with AGNs and starburst galaxies, which form stars at exceptionally high rates. However, Hayes and colleagues' discovery[1] of polarized emission from the largest and brightest of the blobs now argues strongly in favour of the second option.

To appreciate the significance of this result, it is essential to understand what makes the emission from the blobs special. The emission is seen only at the wavelength corresponding to the Lyman-α electronic transition of hydrogen. The Lyman-α spectral line is created by a transition from the first excited state of hydrogen to the ground state. Because atoms in the ground state are common, the transition is resonant, meaning that a Lyman-α photon that is emitted by one atom is quickly absorbed by a neighbouring atom, promoting it to the excited state. This atom quickly decays to the ground state, releasing another Lyman-α photon, which will be absorbed by another atom, and so on.

It is therefore difficult for a Lyman-α photon to escape unless the medium is very highly ionized or the random motions of the atoms cause the photon to undergo Doppler shifting to a higher or lower frequency, so that the frequency of the photon no longer matches the Lyman-α transition of a hydrogen atom. Although this scattering in frequency leads to the photon's escape, it also leaves an imprint on the polarization of the photon because escaping photons are, on average, scattered perpendicular to their initial direction of travel[7]. A similar process is responsible for polarizing the light reflected from the surface of water, and is why many sunglasses contain polarized lenses. Reducing this complex problem to the simplest level, rings of polarized emission would be expected to be observed if the photons originate from a central source, whereas unpolarized emission would be expected if the source of the photons is widely distributed.

In Hayes and colleagues' technically impressive study[1] of the emission from the brightest of the Lyman-α blobs, the polarization pattern consists of concentric rings around the likely centre of the emission (Fig. 1), in good agreement with the predictions for scattering from a central source. If this result can be generalized to all Lyman-α blobs, it opens up the intriguing possibility of using the properties of the scattered radiation to probe the gas haloes of young galaxies. When combined with absorption measurements from background galaxies and AGNs, observations of extended Lyman-α blobs suggest a promising route to studying the previously invisible extended circum-galactic medium of galaxies[8]. This possibility is exciting, because stars and cold gas account for only a small fraction of the ordinary matter in the Universe: theoretical models predict that most ordinary matter will reside in an extended and hitherto undetectable reservoir. ■

**Richard Bower** *is at the Institute for Computational Cosmology, Department of Physics, University of Durham, Durham DH1 3LE, UK.*
*e-mail: r.g.bower@durham.ac.uk*

1. Hayes, M., Scarlata, C. & Siana, B. *Nature* **476,** 304–307 (2011).
2. Steidel, C. C. *et al. Astrophys. J.* **532,** 170–182 (2000).
3. Fardal, M. A. *et al. Astrophys. J.* **562,** 605–617 (2001).
4. Goerdt, T. *et al. Mon. Not. R. Astron. Soc.* **407,** 613–631 (2010).
5. Matsuda, Y. *et al. Astron. J.* **128,** 569–584 (2004).
6. Geach, J. E. *et al. Astrophys. J.* **700,** 1–9 (2009).
7. Dijkstra, M. & Loeb, A. *Mon. Not. R. Astron. Soc.* **386,** 492–504 (2008).
8. Steidel, C. C. *et al. Astrophys. J.* **736,** 160 (2011).



© 2011 Macmillan Publishers Limited. All rights reserved

GLACIOLOGY

# Past ice-shelf collapse in West Antarctica

**Images of the sea floor in Pine Island Bay, West Antarctica, reveal impressive evidence that a massive ice-shelf break-up occurred before about 12,000 years ago, and point to a tidal influence on sea-floor features produced during deglaciation.**

COLM Ó COFAIGH

Recent disintegration of the floating ice shelves that fringe the Antarctic Peninsula[1,2] has provided dramatic evidence of the effects of climate change (Fig. 1). Glaciologists have long regarded such ice shelves as critical to the stability of the Antarctic ice sheet, because they act as a buffer between the grounded, often fast-flowing, parts of the ice sheet and the surrounding ocean, in effect providing a control on the discharge of grounded ice. In the Antarctic Peninsula, satellite-derived observations of glacier velocity have shown that glaciers there accelerated after ice-shelf collapse[3]. Writing in *Geology*, Jakobsson *et al.*[4] now present sea-floor images showing that ice-shelf break-up also occurred in Pine Island Bay in West Antarctica during the ice-sheet retreat that occurred before about 12,000 years ago.

This finding[4] is notable because it indicates that the larger West Antarctic ice shelves have also collapsed in the past, in this case in an area of the ice sheet that is currently responsible for almost 30% of ice drainage from West Antarctica and that has seen recent thinning and accelerated flow[5]. It is also noteworthy because Jakobsson *et al.* propose that the icebergs produced during this collapse, and the associated sea-floor features, were strongly influenced by ocean tides. Although such features have been imaged previously in several areas of the Antarctic continental shelf, the new data[4] show them in unprecedented detail.

Using geophysical equipment mounted in the hull of an icebreaking research ship, Jakobsson and colleagues carried out a detailed survey of a 4,140-square-kilometre area of the sea floor of Pine Island Bay. Here, a bathymetric trough — Pine Island Trough — extends through the bay, and at the Last Glacial Maximum about 18,000 years ago, it contained a grounded, and probably fast-flowing, glacier. Jakobsson and colleagues' data reveal surface features in the trough that were produced by icebergs following the break-up of an ice shelf in Pine Island Bay. These landforms comprise linear to curvilinear sets of furrows, oriented along the trough, that are eroded into a relatively flat area of the sea floor and are spaced 150–500 metres apart. Within these furrows, and oriented perpendicular to them, are beautifully preserved, very regular small ridges 1–2 m high, spaced 60–200 m apart. The ridges have a corrugated appearance on the seabed, and bends or kinks are repeated from one ridge to the next. Overall ridge spacing decreases progressively seawards. The authors[4] interpret both the furrows and ridges as a product of the action of icebergs produced during a massive ice-shelf breakout event.

Jakobsson *et al.* propose that the furrows were produced by icebergs grounded on the sea floor and dragged across it, probably owing to the action of wind and/or currents, resulting in the production of a series of longitudinal scours. The small corrugated ridges are remarkable for the regularity of their spacing, and the authors argue that this implies that a periodic forcing was in operation during their formation. They propose that this forcing was a diurnal tidal cycle, and that the ridges originated from the lifting and settling of 'mega-icebergs' on the sea floor in rhythm with the tide.

From the well-preserved nature of the ridges and their relatively uniform spacing and length, Jakobsson *et al.* infer that the mega-icebergs formed part of a much larger, coherent mass of ice that broke off at the grounding line (the transition from grounded ice to a floating ice shelf) and remained upright and largely intact following break-off. Each time the mass of icebergs settled on the sea floor, soft sediment was squeezed up behind, forming a ridge. A series of ridges was therefore produced in the wake of the drifting mass of icebergs before its eventual grounding and fragmentation in shallower water.

What caused the disintegration of the Pine Island Bay ice shelf, and what effect did this have on the grounding line? From dating marine sediment cores, Jakobsson *et al.*[4] propose that breakout of the Pine Island Bay ice shelf occurred before about 12,000 years ago and that ice-shelf break-up may have been driven by some combination of sea-level rise and ocean warming. They further infer that ice-shelf break-up resulted in the retreat of the grounding line in Pine Island Bay.

Conceptually, this inference is reasonable by analogy with recent evidence from the Antarctic Peninsula[3]. However, although there is indeed support for recession of grounded ice in Pine Island Bay[6,7], direct evidence for a causal connection between the ice-shelf disintegration documented by Jakobsson and colleagues and grounding-line recession in the bay remains inferential and requires further dating control to demonstrate the link conclusively. Admittedly, the acquisition of such dating control is a challenge in Antarctica, where carbonate material, which can be radiocarbon dated to constrain the period of deglaciation, is often sparse in marine sediments.



**Figure 1 | Ice-shelf break-up.** The image shows icebergs from the Larsen B ice shelf, Antarctic Peninsula, after its collapse in 2002. Jakobsson *et al.*[4] present geophysical evidence from the sea floor of Pine Island Bay, West Antarctica, for a similar large-scale ice-shelf break-up during the last deglaciation. (Photo taken March 2002 by C.ÓC.)



© 2011 Macmillan Publishers Limited. All rights reserved

The authors' preferred model, in which the icebergs remained intact and upright after initial breakout, also contrasts with the modern Larsen B ice shelf, where icebergs capsized during collapse[8].

It is particularly noteworthy that, as Jakobsson *et al.* emphasize, the corrugated ridges and furrows on the floor of Pine Island Bay occur widely on the Antarctic continental shelf, including in the Ross and Weddell seas. This suggests that tides may be a more pervasive influence on the record of Antarctic ice-sheet deglaciation than has hitherto been recognized, and it raises the interesting question of whether spatial and temporal variability in tidal range could act as a control on ice-sheet retreat[9]. Clearly, much remains to be discovered about the complexity of the geological record of West Antarctic Ice Sheet deglaciation and the associated controls. ■

**Colm Ó Cofaigh** *is in the Department of Geography, Durham University, Durham DH1 3LE, UK.*
*e-mail: colm.ocofaigh@durham.ac.uk*

1. Vaughan, D. G. & Doake, C. S. M. *Nature* **379,** 328–331 (1996).
2. Domack, E. *et al. Nature* **436,** 681–685 (2005).
3. De Angelis, H. & Skvarca, P. *Science* **299,** 1560–1562 (2003).
4. Jakobsson, M. *et al. Geology* **39,** 691–694 (2011).
5. Wingham, D. J., Wallis, D. W. & Shepherd, A. *Geophys. Res. Lett.* **36,** L17501; http://dx.doi.org/10.1029/2009GL039126 (2009).
6. Lowe, A. L. & Anderson, J. B. *Quat. Sci. Rev.* **21,** 1879–1897 (2002).
7. Graham, A. G. C. *et al. J. Geophys. Res.* **115,** F03025; http://dx.doi.org/10.1029/2009JF001482 (2010).
8. MacAyeal, D. R., Scambos, T. A., Hulbe, C. L. & Fahnestock, M. A. *J. Glaciol.* **49,** 22–36 (2003).
9. Uehara, K., Scourse, J. D., Horsburgh, K. J., Lambeck, K. & Purcell, A. P. *J. Geophys. Res.* **111,** C09025; http://dx.doi.org/10.1029/2006JC003531 (2006).

LINGUISTICS

# Deep relationships between languages

**Tracing a common ancestry between languages becomes harder as the connection goes further back in time. A new test has revealed a surprisingly ancient relationship between a central Siberian and a North American language family.**

JARED DIAMOND

Humans crossed from Siberia into Alaska some 14,000 years ago, long before the 5,000–10,000-year barrier beyond which most linguists think that language evolution erases traces of shared ancestry. Efforts to break this barrier by establishing a relationship between Old World and New World languages have been unconvincing until now. The change is due mainly to the efforts of Edward Vajda, who offers evidence in *Anthropological Papers of the University of Alaska*[1] for the first convincing link, over an astonishingly large distance, between Old World and New World languages.

Most of the world's 7,000 present-day languages fall into language families — such as the Indo-European family, to which this journal's English, along with 140 other languages, belongs. Languages of the same family resemble each other because they arose from a common ancestor and then diverged.

The difficulty in detecting ancient linguistic relationships — those beyond the 5,000–10,000-year barrier — is that all languages share the same pool of consonants, vowels and grammatical structures. Therefore, some resemblances are expected to occur by chance if one compares lists of thousands of words for two unrelated languages. For instance, the word for 'eel' is almost the same in Japanese and in New Guinea's Haruai language, despite a lack of other similarities between the two languages and the impossibility that one could have borrowed the word from the other[1].

To address this problem, linguists scrutinize word lists for putative cognates — phonetically similar words with the same meaning in two related languages, and differing by regular sound correspondences — such as English and Spanish cognates beginning with 'f' and 'p', respectively (father, fish and foot versus padre, pez and pie). The late Joseph Greenberg emphasized the value of multilateral comparisons, in which many related languages are simultaneously compared to identify ancestral roots preserved in only some of the daughter languages. He concluded that most Native American languages belong to one large ancient family, termed Amerind[2], and that Indo-European and other established families form a superfamily, termed Nostratic. Twenty-seven shared worldwide etymologies suggested even more ancient shared ancestry among all the world's languages[3]. However, most linguists reject these conclusions, in part because of a lack of statistical tests to prove that the claimed similarities are not due to chance.

There are two methods for testing the role




## 50 Years Ago

On the whole the average alcohol consumption within a species of animal is constant. It can be said that each species has its own specific liking for alcohol … In this work the consumption of 10 per cent alcohol and water in free-choice experiments in six animal species has been studied … The hedgehogs preferred water to alcohol solution. Sometimes, however, the animals drank considerable amounts of alcohol, and it could at times be judged from their movements that they were slightly drunk. The hamsters preferred alcohol. They seldom tasted the water and, when they did so, only until they found out that they were drinking from the wrong bottle.
**From *Nature* 19 August 1961**

## 100 Years Ago

On the publication of the final report of the Royal Commission on Tuberculosis the view was frequently expressed that those in authority ought, long ago, to have taken precautions against the dangers arising out of the use of milk containing tubercle bacilli … Now that the commissioners have reported, and in no uncertain voice, that tuberculosis, especially in the child, may be the result of infection with tubercle bacilli conveyed in the cow's milk, it is essential that the question of regulations relating to milk and meat supply should be carefully reconsidered, and that, as the commissioners put it, "Government should cause to be enforced throughout the kingdom food regulations, planned to afford better security against the infection of human beings through the medium of diet derived from tuberculous animals."
**From *Nature* 17 August 1911**


© 2011 Macmillan Publishers Limited. All rights reserved

The authors' preferred model, in which the icebergs remained intact and upright after initial breakout, also contrasts with the modern Larsen B ice shelf, where icebergs capsized during collapse[8].

It is particularly noteworthy that, as Jakobsson *et al.* emphasize, the corrugated ridges and furrows on the floor of Pine Island Bay occur widely on the Antarctic continental shelf, including in the Ross and Weddell seas. This suggests that tides may be a more pervasive influence on the record of Antarctic ice-sheet deglaciation than has hitherto been recognized, and it raises the interesting question of whether spatial and temporal variability in tidal range could act as a control on ice-sheet retreat[9]. Clearly, much remains to be discovered about the complexity of the geological record of West Antarctic Ice Sheet deglaciation and the associated controls. ■

**Colm Ó Cofaigh** *is in the Department of Geography, Durham University, Durham DH1 3LE, UK.*
*e-mail: colm.ocofaigh@durham.ac.uk*

1. Vaughan, D. G. & Doake, C. S. M. *Nature* **379,** 328–331 (1996).
2. Domack, E. *et al. Nature* **436,** 681–685 (2005).
3. De Angelis, H. & Skvarca, P. *Science* **299,** 1560–1562 (2003).
4. Jakobsson, M. *et al. Geology* **39,** 691–694 (2011).
5. Wingham, D. J., Wallis, D. W. & Shepherd, A. *Geophys. Res. Lett.* **36,** L17501; http://dx.doi.org/10.1029/2009GL039126 (2009).
6. Lowe, A. L. & Anderson, J. B. *Quat. Sci. Rev.* **21,** 1879–1897 (2002).
7. Graham, A. G. C. *et al. J. Geophys. Res.* **115,** F03025; http://dx.doi.org/10.1029/2009JF001482 (2010).
8. MacAyeal, D. R., Scambos, T. A., Hulbe, C. L. & Fahnestock, M. A. *J. Glaciol.* **49,** 22–36 (2003).
9. Uehara, K., Scourse, J. D., Horsburgh, K. J., Lambeck, K. & Purcell, A. P. *J. Geophys. Res.* **111,** C09025; http://dx.doi.org/10.1029/2006JC003531 (2006).

LINGUISTICS

# Deep relationships between languages

**Tracing a common ancestry between languages becomes harder as the connection goes further back in time. A new test has revealed a surprisingly ancient relationship between a central Siberian and a North American language family.**

JARED DIAMOND

Humans crossed from Siberia into Alaska some 14,000 years ago, long before the 5,000–10,000-year barrier beyond which most linguists think that language evolution erases traces of shared ancestry. Efforts to break this barrier by establishing a relationship between Old World and New World languages have been unconvincing until now. The change is due mainly to the efforts of Edward Vajda, who offers evidence in *Anthropological Papers of the University of Alaska*[1] for the first convincing link, over an astonishingly large distance, between Old World and New World languages.

Most of the world's 7,000 present-day languages fall into language families — such as the Indo-European family, to which this journal's English, along with 140 other languages, belongs. Languages of the same family resemble each other because they arose from a common ancestor and then diverged.

The difficulty in detecting ancient linguistic relationships — those beyond the 5,000–10,000-year barrier — is that all languages share the same pool of consonants, vowels and grammatical structures. Therefore, some resemblances are expected to occur by chance if one compares lists of thousands of words for two unrelated languages. For instance, the word for 'eel' is almost the same in Japanese and in New Guinea's Haruai language, despite a lack of other similarities between the two languages and the impossibility that one could have borrowed the word from the other[1].

To address this problem, linguists scrutinize word lists for putative cognates — phonetically similar words with the same meaning in two related languages, and differing by regular sound correspondences — such as English and Spanish cognates beginning with 'f' and 'p', respectively (father, fish and foot versus padre, pez and pie). The late Joseph Greenberg emphasized the value of multilateral comparisons, in which many related languages are simultaneously compared to identify ancestral roots preserved in only some of the daughter languages. He concluded that most Native American languages belong to one large ancient family, termed Amerind[2], and that Indo-European and other established families form a superfamily, termed Nostratic. Twenty-seven shared worldwide etymologies suggested even more ancient shared ancestry among all the world's languages[3]. However, most linguists reject these conclusions, in part because of a lack of statistical tests to prove that the claimed similarities are not due to chance.

There are two methods for testing the role




## 50 Years Ago

On the whole the average alcohol consumption within a species of animal is constant. It can be said that each species has its own specific liking for alcohol … In this work the consumption of 10 per cent alcohol and water in free-choice experiments in six animal species has been studied … The hedgehogs preferred water to alcohol solution. Sometimes, however, the animals drank considerable amounts of alcohol, and it could at times be judged from their movements that they were slightly drunk. The hamsters preferred alcohol. They seldom tasted the water and, when they did so, only until they found out that they were drinking from the wrong bottle.
**From *Nature* 19 August 1961**

## 100 Years Ago

On the publication of the final report of the Royal Commission on Tuberculosis the view was frequently expressed that those in authority ought, long ago, to have taken precautions against the dangers arising out of the use of milk containing tubercle bacilli … Now that the commissioners have reported, and in no uncertain voice, that tuberculosis, especially in the child, may be the result of infection with tubercle bacilli conveyed in the cow's milk, it is essential that the question of regulations relating to milk and meat supply should be carefully reconsidered, and that, as the commissioners put it, "Government should cause to be enforced throughout the kingdom food regulations, planned to afford better security against the infection of human beings through the medium of diet derived from tuberculous animals."
**From *Nature* 17 August 1911**


© 2011 Macmillan Publishers Limited. All rights reserved




**Figure 1 | Circumpolar map of the Arctic showing the locations of Yeniseian and Na-Dene languages.** There are two small groups of outlying Na-Dene populations in North America, far south of those depicted, in the United States southwest and on the California and Oregon coasts. Vajda[1] has found evidence of a link between the Yeniseian and Na-Dene languages. (Figure based on ref. 1.)

of chance in assessing deep relationships[1,4–7]. One is also used by biologists to test evolutionary trees: the Monte Carlo simulation. This involves taking word lists for two or more languages, repeatedly and randomly scrambling the meanings within each list, recounting the number of cognates in each randomized pair of lists, and then calculating statistically whether the observed number of cognates exceeds that expected by chance, according to the randomized lists.

Such tests have been extended by Kessler and Lehtonen[4] to multilateral comparisons, after first confirming statistical significance among 11 languages already known to belong to the Indo-European family. This was true even for Albanian, a language whose Indo–European affinities had proved difficult to establish by bilateral comparisons. This success vindicates Greenberg's view that multilateral comparisons can uncover evidence of a relationship that is obscure in bilateral comparisons (because any single Indo-European language alone happens to lack certain Indo-European roots retained in Albanian).

Unhappily for those multilateralists who consider Indo-European to have an ancient relationship with Uralic languages (which include Finnish and Hungarian), the two families turned out to be related at a significance level of only $P = 0.45$, far from the $P < 0.05$ required for significance[4]. It seems likely that the few sharings between the two families are due to chance or to borrowing.

The other, more approximate, set of methods for assessing the role of chance depends on 'individual-identifying thresholds'[1,5,6]. With about 7,000 languages in the world, one could estimate $P = 1/7{,}000 = 0.000143$ as a level of rarity to suggest a unique individual language, or $(1/7{,}000)\,(0.05) = 7 \times 10^{-6}$ as a level of rarity significant at $P < 0.05$. Johanna Nichols[1] uses this test to support Vajda's remarkable new evidence for a relationship between central Siberian Yeniseian languages and North American Na-Dene languages.

Yeniseian is a tiny family of less than a dozen languages that were formerly spoken along Siberia's Yenisei River, and now survive only as the Ket language, which has fewer than 200 speakers. Na-Dene languages (that is, the Tlingit, Eyak and Athabaskan languages) are spoken across much of Alaska and northwest Canada, with outliers in the southwestern United States, California and Oregon. Four intervening language families and 5,200 kilometres separate the nearest Ket and Na-Dene speakers today (Fig. 1).

In 1923, one possible cognate was found between Ket, Athabaskan and Tlingit, and other similar words were later identified (see ref. 3 for examples). Vajda has now placed this postulated relationship on a rigorous footing. He was struck by how Ket's system of tones and bewilderingly complex strings of eight verb prefixes were utterly out of place in Siberia — otherwise occupied by toneless suffixing languages related to Turkish and Finnish — and by how those tones and prefixes corresponded in detail to Na-Dene languages. The parallels he identifies include a dozen grammatical prefixes and about 100 cognate words with sound correspondences. Because the prefixes appear in the same sequence between Yeniseian and Na-Dene verbs, and because Vajda analysed all prefixes and didn't just fish for one matching pair, both his prefix comparisons and his word lists comfortably pass Nichols' significance tests. Vajda's evidence is judged to be plausible by linguists previously sceptical of claimed relationships between language families[1].

The reported Yeniseian–Na-Dene link raises many questions. Where was the common ancestral language spoken (Siberia or North America), and did Yeniseians move west or Na-Denes move east across the Bering Strait? Are there archaeological correlates of such a movement? (The only plausible candidates are archaeological horizons at around 12,000 and 5,000 years ago.) Do Yeniseian and Na-Dene form a superfamily by themselves, or do they belong to a larger superfamily that includes Sino-Tibetan and Basque? Are Yeniseian and Na-Dene people more closely related genetically to each other than to other Siberians and Native Americans? (Apparently not, possibly because of millennia of intermarriage with surrounding peoples.) Why did Yeniseian speakers have such a tiny geographical range in modern times?

But there are two questions that most trouble linguists. Why do Yeniseian and Na-Dene languages still show such a strong relationship if they diverged 12,000 years ago, when other languages diverge beyond recognition after 5,000–10,000 years ago? Either Yeniseian and Na-Dene languages really diverged only 5,000 years ago, or they are unusually conservative and evolve especially slowly. And how did Yeniseians and Na-Denes get separated by 5,200 km in only 5,000 or 12,000 years, making their overland migration the longest and fastest recorded by hunter-gatherers? In coming years, we can watch as these questions are debated — and perhaps as other links between Old and New World language families are revealed. ■

**Jared Diamond** *is in the Geography Department, University of California, Los Angeles, California 90095-1524, USA.*
*e-mail: jdiamond@geog.ucla.edu*

1. Kari, J. & Potter, B. (eds) *Anthropological Papers of the University of Alaska* Vol. 5 (issues 1, 2) (Univ. Alaska Fairbanks/Alaska Native Language Center, 2010).
2. Greenberg, J. H. *Language in the Americas* (Stanford Univ. Press, 1987).
3. Ruhlen, M. *On the Origin of Languages* (Stanford Univ. Press, 1994).
4. Kessler, B. & Lehtonen, A. in *Phylogenetic Methods and the Prehistory of Languages* (eds Forster, P. & Renfrew, C.) 33–42 (McDonald Institute, 2006).
5. Nichols, J. in *The Handbook of Language Contact* (ed. Hickey, R.) 361–379 (Wiley-Blackwell, 2010).
6. Nichols, J. & Warnow, T. *Lang. Ling. Compass* **2,** 760–820 (2008).
7. Wichmann, S. & Holman, E. W. *Assessing Temporal Stability for Linguistic Typological Features* (Lincom Europa, 2009).


© 2011 Macmillan Publishers Limited. All rights reserved

ARTICLE

doi:10.1038/nature10337

# Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling

Wim de Lau[1]*, Nick Barker[1]†*, Teck Y. Low[2], Bon-Kyoung Koo[1], Vivian S. W. Li[1], Hans Teunissen[1], Pekka Kujala[3], Andrea Haegebarth[1]†, Peter J. Peters[3], Marc van de Wetering[1], Daniel E. Stange[1], Johan E. van Es[1], Daniele Guardavaccaro[1], Richard B. M. Schasfoort[4], Yasuaki Mohri[5], Katsuhiko Nishimori[5], Shabaz Mohammed[2], Albert J. R. Heck[2] & Hans Clevers[1]

**The adult stem cell marker Lgr5 and its relative Lgr4 are often co-expressed in Wnt-driven proliferative compartments. We find that conditional deletion of both genes in the mouse gut impairs Wnt target gene expression and results in the rapid demise of intestinal crypts, thus phenocopying Wnt pathway inhibition. Mass spectrometry demonstrates that Lgr4 and Lgr5 associate with the Frizzled/Lrp Wnt receptor complex. Each of the four R-spondins, secreted Wnt pathway agonists, can bind to Lgr4, -5 and -6. In HEK293 cells, RSPO1 enhances canonical WNT signals initiated by WNT3A. Removal of LGR4 does not affect WNT3A signalling, but abrogates the RSPO1-mediated signal enhancement, a phenomenon rescued by re-expression of LGR4, -5 or -6. Genetic deletion of *Lgr4/5* in mouse intestinal crypt cultures phenocopies withdrawal of Rspo1 and can be rescued by Wnt pathway activation. Lgr5 homologues are facultative Wnt receptor components that mediate Wnt signal enhancement by soluble R-spondin proteins. These results will guide future studies towards the application of R-spondins for regenerative purposes of tissues expressing Lgr5 homologues.**

The genes *Lgr4*, *Lgr5* and *Lgr6* encode orphan 7-transmembrane receptors that are close relatives of the receptors for follicle stimulating hormone (FSH), luteinizing hormone (LH) and thyroid-stimulating hormone (TSH)[1]. It is currently unknown how they signal and what their ligands are. *Lgr5* is a Wnt target gene that marks proliferative stem cells in several Wnt-dependent stem cell compartments, that is, the small intestine and colon[2], the stomach[3] and the hair follicle[4]. *Lgr6* marks multipotent stem cells in the epidermis[5]. The expression of *Lgr4* is much broader[6], but we noted that *Lgr5* is co-expressed with *Lgr4* in the stem cell compartments mentioned above. For instance, *Lgr5* marks small intestinal stem cells at the crypt base, whereas *Lgr4* marks all crypt cells, including $Lgr5^+$ stem cells (Fig. 1a). The same observation was made in ref. 7. Indeed, our previously reported microarray study, which compared $Lgr5^+$ stem cells to their immediate daughters, showed that *Lgr4* is expressed in both cell types[8]. Moreover, we generated an allele of *Lgr4* in which a green fluorescent protein-internal ribosomal entry site-CreERT2 (GFP-IRES-CreERT2) cassette was inserted into the ATG start codon, as performed previously for *Lgr5* (ref. 2) and *Lgr6* (ref. 5). Tamoxifen-induced lineage tracing after crossing to the R26R-LacZ Cre reporter strain demonstrated that *Lgr4* is expressed both by short-lived progenitors and by long-lived stem cells (Supplementary Fig. 1). Of note, *Lgr4* and *Lgr5* null mutations are neonatal-lethal in mice[9,10].

To address a potential function in crypts, we generated an $Lgr5^{fl}$ allele (Supplementary Fig. 2a) by flanking exon 16 with *loxp* sites; its deletion causes a frame shift. This allele was crossed into a mouse strain carrying conditional $Lgr4^{fx}$ alleles[10] and the gut-specific *Ah-Cre* transgene, which is inducible by β-naphtoflavone[11]. Conditional deletion of *Lgr5* alone in the intestinal epithelium of adult mice yielded no apparent phenotype, and did not confirm the Paneth cell phenotype reported previously in *Lgr5* null neonatal mice[12]. Deletion of *Lgr4* alone induced loss of proliferation and crypt loss, obvious from day 4–5 post-induction onwards. Paneth cells at crypt bottoms were retained in 'nests', disconnected from the surface epithelium. No direct effects were observed on differentiated cells. The combined deletion of *Lgr4* and *-5* enhanced this crypt phenotype as judged by the cell proliferation marker Ki67 and the stem cell marker olfactomedin 4 (Olfm4; ref. 13). Figure 1b and Supplementary Fig. 2b depict typical results obtained at day 5 post-induction. Of note, hyperplastic wild-type 'escaper' crypts (arrows in Fig. 1b, and see ref. 14) served as staining control. Over the next few days, villi shortened and eventually the phenotype was not compatible with life. A gut phenotypic analysis was reported recently for mice homozygous for a hypomorphic *Lgr4* allele that live for about 1 month postnatally[7]. These mice displayed twofold reduced proliferation and an 85% reduction in Paneth cells, whereas stem cell markers and Wnt target genes were unaffected *in vivo*. In 'minigut' culture[15], *Lgr4*-hypomorphic crypts failed to initiate organoid growth. No genetic interaction with an *Lgr5* null allele was observed for this intestinal phenotype.

Two signalling pathways, Wnt[16,17] and Notch[18], are crucial for the maintenance of adult crypt proliferation. We determined gene expression changes by microarraying on day 1 post-deletion of *Lgr4* and *-5*, when stem cells were physically still present (Supplementary Fig. 3). Simultaneous deletion of *Lgr4* and *Lgr5* resulted in downregulation of 306 genes in two separate experiments (*Lgr4/5* gene set; Supplementary Table 1). These included 29 stem cell-enriched genes[8]. To determine if these are Wnt target genes, we analysed this gene set in two complementary scenarios that detect intestinal Wnt target genes. In the first scenario[19], *in vivo* deletion of *Apc* results in the immediate upregulation of Wnt target genes. Comparison of the *Lgr4/5* gene set to the microarray data from ref. 19 showed a significant upregulation of 53% (137 genes) of the *Lgr4/5* gene set, whereas only 5% (12 genes) were significantly downregulated (Fig. 2a, red ratios). Indeed, gene set enrichment analysis (GSEA)[20] showed a highly significant enrichment (false

[1]Hubrecht Institute and University Medical Center Utrecht, 3584 CX Utrecht, The Netherlands. [2]Biomolecular Mass Spectrometry and Proteomics group, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University. Netherlands Proteomics Centre, Padualaan 8, 3584 CH Utrecht, The Netherlands. [3]Antoni van Leeuwenhoek Hospital Netherlands Cancer Institute, 1066 CX Amsterdam, The Netherlands. [4]Medical Cell Biophysics, MIRA institute, University of Twente, 7500 AE Enschede, The Netherlands. [5]Tohoku University, 981-8555 Sendai, Miyagi, Japan. †Present addresses: Institute of Medical Biology, 06-06 Immunos, Singapore 138648 (N.B.); Bayer Schering Pharma AG, 13353 Berlin, Germany (A.H.).
*These authors contributed equally to this work.



©2011 Macmillan Publishers Limited. All rights reserved

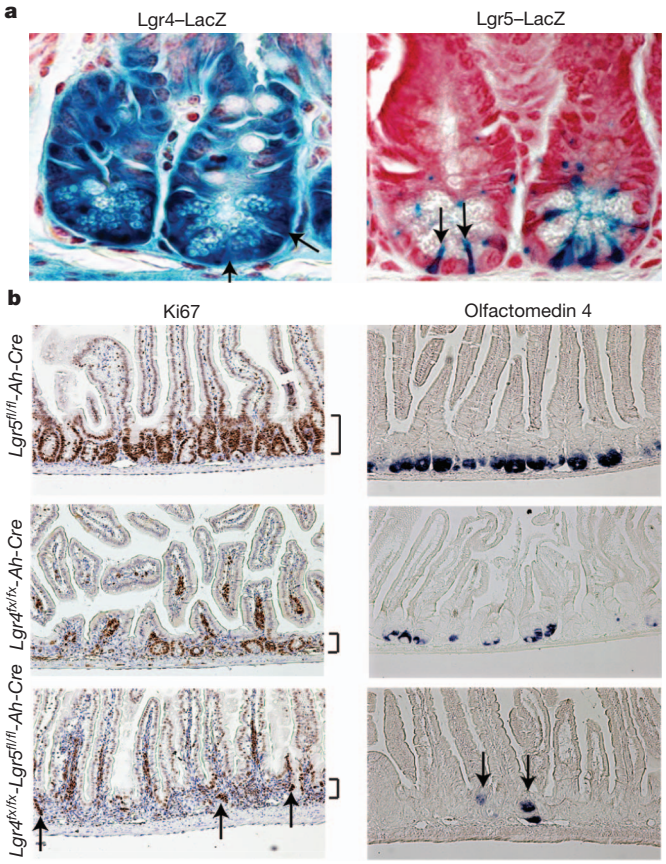



**Figure 1 | Conditional deletion of *Lgr4* and *Lgr5*.** **a**, Left, expression of Lgr4–lacZ[6] (blue) occurs throughout intestinal crypts. Right, expression of Lgr5–LacZ (blue) is specific to stem cells (arrows) located between Paneth cells at bottom of crypts[2]. **b**, Adult $Lgr4^{fx/fx}$ and/or $Lgr5^{fl/fl}$ mice carrying the *Ah-Cre* transgene, analysed 5 days after Cre-activation[11]. Left panels, proliferation visualized by Ki67. Vertical bars, crypt width. Right panels, stem cells visualized by *OlfM4 in situ* hybridization. Top, *Lgr5* deletion has no obvious effect. Middle, *Lgr4* deletion has significant deleterious effects on crypt stem cells and proliferative progenitors. Bottom, upon *Lgr4/5* double deletion, >80% of crypts entirely disappear. Arrows indicate wild-type escaper crypts that express OlfM4 and are $Ki67^{+}$.

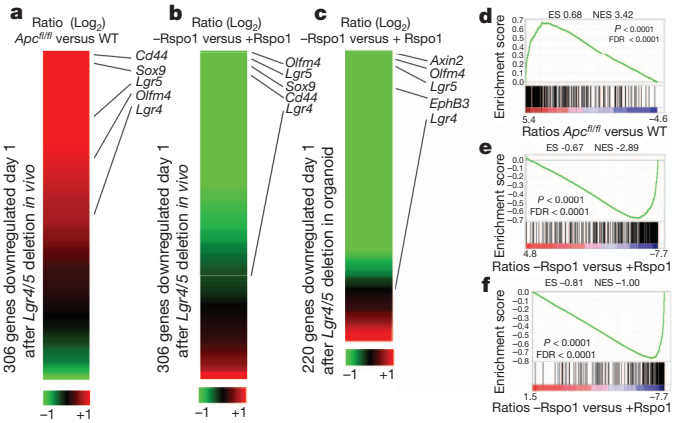



**Figure 2 | Wnt target genes are downstream of Lgr4/5.** Concomitant deletion of *Lgr4* and *Lgr5* in $Lgr4^{fl/fl}$ $Lgr5^{fl/fl}$ mice resulted in downregulation of 307 unique genes 1 day after deletion (Supplementary Table 1). The intestinal Wnt signature can be revealed by two opposing experiments: deletion of *Apc in vivo* results in upregulation of a Wnt target signature[19], whereas Rspo1 withdrawal from intestinal organoids[15] results in downregulation of this signature. **a**, Heat map of the $\log_2$ ratio of $Apc^{fl/fl}$ mice versus control wild-type mice (WT) for the 307 *Lgr4/5* genes 3 days after deletion of *Apc* (ratios taken from ref. 19). **b**, Heat map of the $\log_2$ ratio of intestinal organoids 1 day after Rspo1 withdrawal (−Rspo1) versus control organoids (+Rspo1) for the 307 *Lgr4/5* gene set. **c**, Heat map of the $\log_2$ ratio of intestinal organoids 1 day after Rspo1 withdrawal (−Rspo1) versus control organoids (+Rspo1) for the 220 downregulated genes in $Lgr4^{fl/fl}$ $Lgr5^{fl/fl}$ organoids (1 day after deletion). **d**, Gene set enrichment analysis (GSEA). Genes are ranked according to their differential expression between *Apc*-deleted and wild-type mice (data from ref. 19). Black bars beneath the graph depict the rank positions of the 306 genes from the *Lgr4/5* gene set. A highly significant enrichment of the 306 *Lgr4/5* genes was detected towards the gene set upregulated 3 days after *Apc* deletion *in vivo*. **e**, Genes ranked according to differential expression in intestinal organoids after Rspo1 withdrawal versus control organoids. GSEA shows a highly significant enrichment of the 306 *Lgr4/5* genes towards the genes downregulated after Rspo1 withdrawal. **f**, Genes are ranked according to their differential expression in intestinal organoids after Rspo1 withdrawal versus control organoids. GSEA shows a highly significant enrichment of the 220 *Lgr4/5* downregulated genes in organoids towards those genes downregulated after Rspo1 withdrawal. ES, enrichment score; FDR, false discovery rate; NES, normalized ES.

discovery rate (FDR) and *P*-value < 0.0001) of the *Lgr4/5* gene set towards the upregulated genes after *Apc* deletion (Fig. 2d). For the second scenario, we performed microarraying before, and 1 day after, acute withdrawal of the Wnt agonist Rspo1 from cultured small intestinal crypt organoids[15]. This resulted in the immediate down-regulation of 38% (166 genes) of the *Lgr4/5* gene set, whereas only 4% (11 genes) showed the opposite behaviour. (Fig. 2b, green ratios). GSEA analysis confirmed the highly significant enrichment (FDR and *P*-value < 0.0001) of the *Lgr4/5* gene set towards the downregulated genes after Rspo1 withdrawal (Fig. 2e).

To find molecular partners of LGR receptors, we pursued a tandem affinity purification mass spectrometric strategy. Bait proteins carried double Flag-haemagglutinin (FH)-tags[21]. First, we transiently transfected tagged Frizzled7 (Frz7–FH; frizzled family receptor 7 is also known as FZD7) into HEK293T cells. We detected significant signatures for Frizzled7, and—as expected—the WNT co-receptors LRP5 and LRP6. Surprisingly, multiple peptides were detected for LGR4 (Supplementary Table 2). Of note, we never observed LGR5 homologues, frizzled or LRP proteins in HEK293 cells using ~20 unrelated baits (Supplementary Table 3). We then generated stable clones of LS174T colorectal cancer cells that moderately overexpress tagged versions of LGR4, LGR5 or Frizzled5. The LGR4 bait captured LGR5, LRP6, Frizzled5 and Frizzled7. The LGR5 bait captured Frizzled6 and LRP5 and LRP6. The Frizzled5 bait captured LGR4, LGR5, LRP5 and LRP6. These proteins were never observed in the non-transfected controls run in parallel. Results are summarized in Supplementary Table 2. Thus, LGR4 and LGR5 can occur in a physical complex with frizzled proteins and LRP5/6.

The four secreted R-spondin proteins activate the canonical Wnt pathway and are particularly potent when synergizing with secreted Wnt proteins[22,23]. Systemic delivery of Rspo1 in mice leads to a dramatic enhancement of the Wnt-responsive intestinal crypts[24] and stimulates epithelial repair[25]. Whereas Rspo3 uses syndecans as receptors to mediate non-canonical Wnt signals[26], the R-spondin receptors that drive canonical Wnt signals have remained controversial[26]. Rspo1 was proposed to bind the Wnt co-receptor Lrp6 (refs 27, 28), to block the Kremen protein that downregulates surface expression of Wnt receptors[29], or to block the interaction of the Wnt inhibitor Dkk1 with Lrp6 (ref. 23). HEK293T cells are responsive to WNT3A and R-spondin[22]. We used this cell line to search for the cognate R-spondin receptor. To validate the efficacy of using secreted baits to detect surface receptors, we generated a tagged version of Dkk1 (Dkk1–FH). Dkk1 interacts with LRP5/6 (ref. 30). We incubated $2 \times 10^9$ HEK293T cells with Dkk1–FH. The cells were washed, lysed and immunoprecipitated with an anti-Flag antibody. The immunoprecipitated material was eluted with Flag peptide, and re-precipitated with anti-haemagglutinin antibody, after which mass spectrometric analysis was performed. We readily identified Dkk1 and endogenous



©2011 Macmillan Publishers Limited. All rights reserved

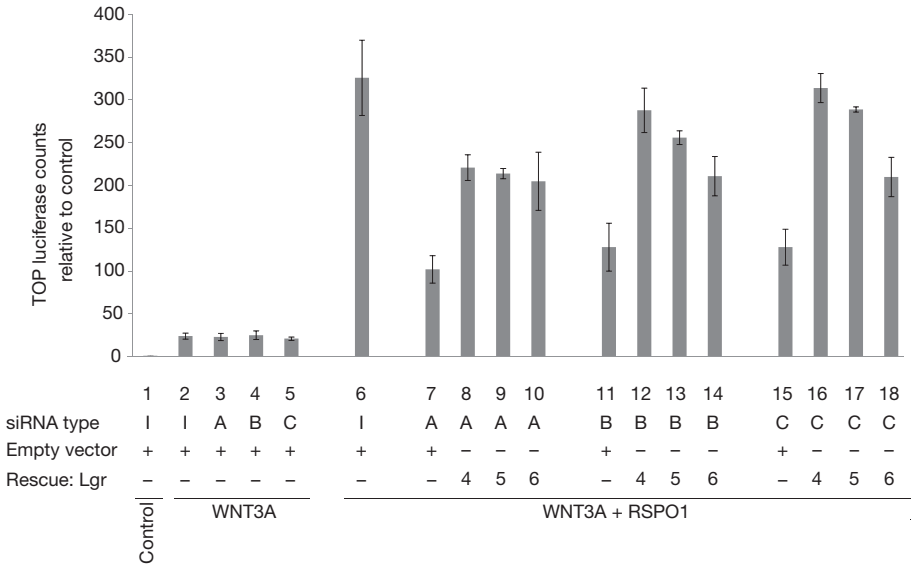



**Figure 3 | LGR4 is essential for transmitting RSPO1 signals but dispensable for transmitting WNT3A signals.** TOPFLASH Wnt reporter assays. HEK293T cells were transfected with indicated siRNAs. siRNAs A, B, C, siRNAs directed at the 3′UTR of human *LGR4*; siRNA I, pool of scrambled siRNA. Three days later, the cells were transfected with TOPFLASH reporters and 5 ng of the indicated rescue constructs, and incubated with WNT3A and RSPO1 as indicated. WNT3A induced TOPFLASH reporter activity without showing effects of removal of LGR4. R-spondin potentiated the WNT3A response, but this effect was sensitive to removal of LGR4. Rescue was obtained with LGR4, -5, and -6 rescue constructs, but not with control vector (lanes 7, 11, 15). $n = 4$. Error bars indicated as s.d.

LRP5 and LRP6 (Supplementary Table 2). We then produced a double-tagged version of human RSPO1 (RSPO1–FH) at ~1 μg ml$^{-1}$. As tested in the TOPFLASH WNT reporter assay[31], the protein potentiated the effect of Wnt3A by approximately 12-fold (Fig. 3). We incubated $2 \times 10^9$ HEK293 cells with the RSPO1–FH-conditioned medium and subjected them to the above protocol. The only transmembrane protein detected was LGR4 (Supplementary Table 2).

HEK293T cells were transiently transfected with tagged versions of LGR4, LGR5 and LGR6, and with the closely related LGR1 (FSH receptor), LGR3 (TSH receptor), LGR7 and LGR8. Transfected cells were incubated with Fc fusion protein of human RSPO1 at ~1 μg ml$^{-1}$, washed, lysed and precipitated with protein G beads. Western blotting for the Flag tag revealed binding of RSPO1–Fc to LGR4, LGR5 and LGR6, but not to the other LGR proteins (Fig. 4a). In the same assay, Fc fusion proteins of RSPO2, RSPO3 and RSPO4 at ~1 μg ml$^{-1}$ were shown to bind specifically to LGR4, LGR5 and LGR6 (Supplementary Fig. 4).

Soluble RSPO1–FH interacted with the leucine-rich-repeat exodomain of LGR5 (amino acids 1 to 546) expressed as a human IgG–Fc fusion protein (LGR5-exo–Fc). This allowed surface plasmon resonance array imaging. Anti-Flag antibody, spotted on the sensor chip, captured RSPO1–FH (Fig. 4b, left, blue line). LGR5-exo–Fc bound to RSPO1–FH (left; blue line). After regeneration, a control experiment was performed using a noggin–Fc fusion protein (Fig. 4b, right, blue line), which did not bind RSPO1–FH. Both LGR5-exo–Fc and noggin–Fc could be captured on anti-human IgG spotted as a control (Fig. 4b, red line). The $K_D$ of the LGR5-exo–Fc–RSPO1–FH interaction was determined at 3.1 nM by an extrapolated 1:1 interaction model at serial injections of RSPO1 and LGR5-exo–Fc on anti-Flag spots with decreasing densities (Supplementary Fig. 5). In a dose–response curve in the HEK293T Wnt reporter assay, Wnt signalling was indeed induced by nanomolar amounts of RSPO1 (Supplementary Fig. 6).

Rat monoclonal antibodies were raised against full-length human LGR5 protein. All antibodies reacted with LGR5-exo–Fc and their

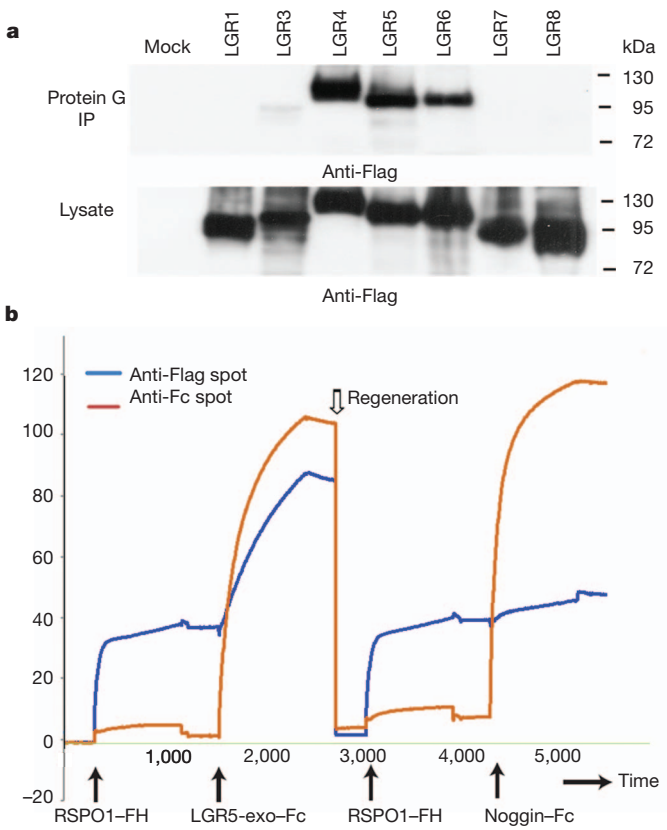



**Figure 4 | Direct physical interaction of RSPO1 with LGR4/5/6 exodomains. a**, RSPO1 binds to LGR4, LGR5 and LGR6. Top panel, HEK293T cells transfected with Flag-tagged versions of the indicated LGR proteins, incubated with RSPO1–Fc fusion protein at ~1 μg ml$^{-1}$. Cells were washed, lysed and RSPO1–Fc was immunoprecipitated with protein G beads. Top, western blotting for Flag revealed specific binding to LGR4, LGR5 and LGR6. IP, immunoprecipitation. Bottom, input of tagged proteins. **b**, RSPO1–LGR5 interaction visualized by surface plasmon resonance. Anti-Flag antibody spotted on the sensor chip captured RSPO1–FH (Fig. 4, left, blue line). After a wash, LGR5-exo–Fc bound to RSPO1–FH. After regeneration with low pH, a noggin–Fc fusion protein served as a negative control (Fig. 3, right, blue line). Both LGR5-exo–Fc and noggin–Fc could be captured on goat anti-human IgG spotted as a control (Fig. 3, red line).



©2011 Macmillan Publishers Limited. All rights reserved

epitopes were mapped using carboxy-terminal deletion clones and human-mouse hybrids of the Lgr5 exodomain (Supplementary Fig. 7a). All antibodies recognizing the amino terminus of the Lgr5 extracellular domain blocked the RSPO1 interaction, whereas the other antibodies did not (Supplementary Fig. 7b). Indeed, when the extreme N terminus or the first LRR domain were deleted from Lgr5, RSPO1-binding was lost (Supplementary Fig. 8).

HEK293T cells express *LGR4* but not *LGR5* (Supplementary Fig. 9). To test if LGR4 constitutes a functional RSPO1 receptor, we removed *LGR4* mRNA from HEK293T cells with three independent small interfering RNAs (siRNAs; Supplementary Fig. 9). Subsequently, the cells were transfected with TOPFLASH WNT reporters and their response to exogenously added WNT3A and RSPO1 was measured. WNT3A alone induced a ~25-fold increase in TOPFLASH activity. Removal of LGR4 had no effect on this WNT3A response (Fig. 3; compare bar 1 to bars 2–5). The combination of WNT3A with RSPO1 led to a further ~12-fold increase in TOPFLASH activity (bar 6). This effect was greatly diminished when LGR4 was removed (bars 7, 11 and 15), but could be rescued by transfection of *LGR4* (bars 8, 12 and 16), *LGR5* (bars 9, 13 and 17) or *LGR6* expression plasmids (bars 10, 14 and 18). No rescue was obtained with *LGR1*, *LGR7* or *LGR8* (Supplementary Fig. 10). A 'DRY' motif at the cytosolic end of the third transmembrane region is believed to be essential for protein G coupling in serpentine receptors[32]. We mutated the ERG motif at the corresponding position in LGR5 to ENG. Both versions of LGR5 rescued loss of LGR4 in the TOPFLASH assay (Supplementary Fig. 11), indicating that protein G signalling is not involved in signalling towards β-catenin.

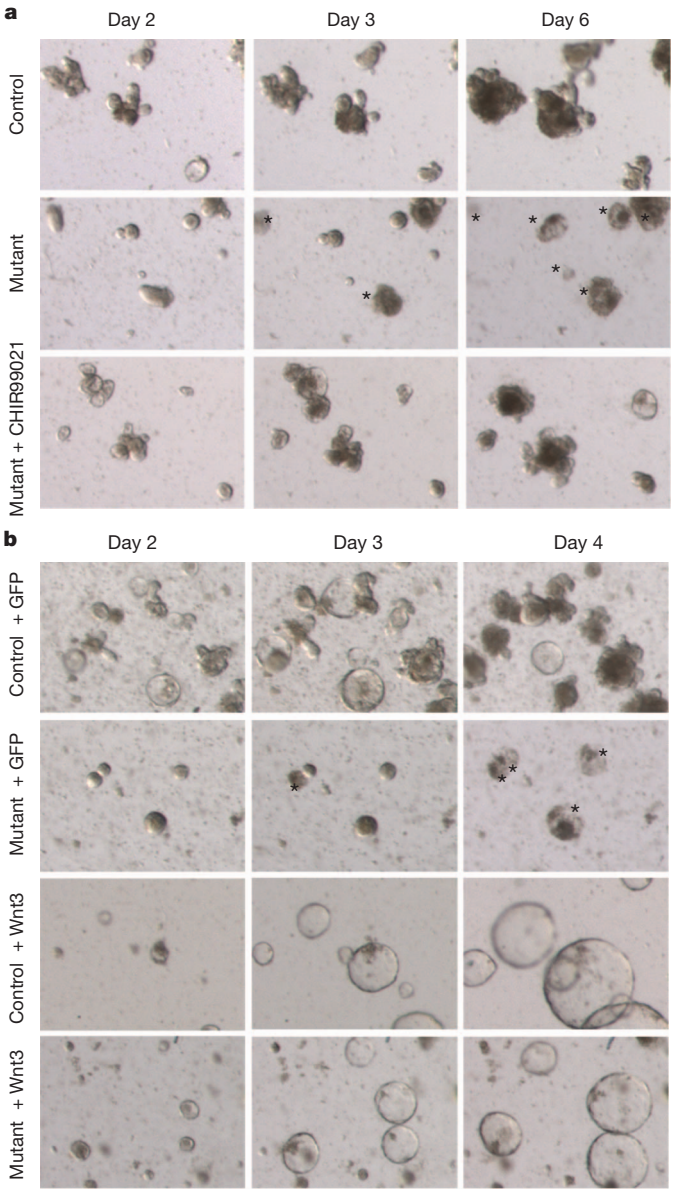



**Figure 5 | Rescue of Lgr4/5 deletion in cultured crypt organoids by Wnt signals. a**, Organoids established under standard conditions[15] from *Lgr4*$^{fx/fx}$ *Lgr5*$^{fl/fl}$ *Villin-CreERT2* mice[37] (Mutant) or from control *Villin-CreERT2* mice (Control). Tamoxifen treatment leads to death (asterisks) of mutant organoids but not that of controls. This is overcome by addition of CHIR99021 at 5 μM (Mutant + CHIR99021). **b**, Organoids established from *Lgr4*$^{fx/fx}$ *Lgr5*$^{fl/fl}$ *Villin-CreERT2* mice (Mutant) or from control *Villin-CreERT2* mice (Control) are infected with Wnt3-expressing retrovirus, which turns organoids into growing, rounded cysts[34], or with control (GFP) retrovirus. Subsequent deletion of *Lgr4* and *-5* has no effect on Wnt3-expressing organoids but leads to death of control retrovirus organoids.

The proposed model implies that strong Wnt signals should rescue loss of Lgr receptors. We generated intestinal organoids[15] from *Lgr4*$^{fx/fx}$ *Lgr5*$^{fl/fl}$ *Villin-CreERT2* mice. Tamoxifen-induced deletion *in vitro* resulted in the demise of the organoids after 2 days (Fig. 5a). Microarraying was performed 1 day after induction of deletion, that is, 24 h before organoids lost viability. This revealed a striking similarity between gene expression changes 1 day after withdrawal of Rspo1 and 1 day after deletion of the proposed R-spondin receptors Lgr4 and Lgr5 (Fig. 2c and f). Addition of the Gsk3 inhibitor CHIR99021 (which activates the Wnt cascade) led to a quantitative, persistent rescue of *Lgr4/5*-mutant organoids (Supplementary Fig. 12). Of note, LiCl (which inhibits Gsk3, but also inositol monophosphatase[34]) reportedly allows 6% rescue of the growth of *Lgr4*-hypomorphic organoids[7]. We next expressed Wnt3 (ref. 35) by retroviral transduction in established *Lgr4*$^{fx/fx}$ *Lgr5*$^{fl/fl}$ *Villin-CreERT2*. Subsequent tamoxifen-induced deletion of *Lgr4* and *-5* led to the demise of the mutant crypts, but this effect was robustly rescued by Wnt3 overexpression (Fig. 5b). In a similar approach, retrovirally expressed Lgr5 and Lgr5-ENG rescued organoid growth (Supplementary Fig. 13).

These data demonstrate that binding of R-spondins to their cognate receptors, the Lgr5 homologues, triggers downstream canonical Wnt signals through associated frizzled–Lrp5/6 complexes. The current observations assign a crucial function to Lgr4 and Lgr5 in Wnt-dependent stem cells and progenitor cells: in intestinal crypts, Lgr4 and -5, incorporated into frizzled–Lrp complexes, allow R-spondins to augment short-range Wnt signals emanating from Paneth cells[34]. This provides a molecular mechanism for the potent hyperplastic response of crypts to Rspo1 (ref. 24) and will guide future studies towards the application of R-spondins for regenerative purposes of tissues expressing Lgr5 homologues. Moreover, this study reinforces the connection, first described in intestinal crypts[36], between Wnt signalling and adult mammalian stem cell biology.

## METHODS SUMMARY

**Mice.** Experiments were performed according to guidelines and reviewed by the DEC of the KNAW. Knock-in alleles were generated and conditional deletion was induced as described elsewhere[2,13].

**Plasmid expression constructs.** Available upon request.

**Histology.** Tissues were prepared and stained with Ki67, PAS, or subjected to fluorescent *in situ* hybridization to *Olfm4* as described elsewhere[13].

**Cell culture, transfections, TOPFLASH assays, RSPO1/WNT3A stimulation.** HEK293T cells and Ls174T cells were cultured and transfected as described elsewhere[31,36]. siRNAs directed at the 3′ untranslated repeat (3′UTR) of human *LGR4* were from Thermo Scientific Dharmacon. Sense sequences: A, 5′-gaaagua aacuguggucaauu-3′; B, 5′-ggguggagucuuaaaguuauu-3′; C, 5′-gguaagaaacuccu aauuauu-3′; I (irrelevant), ON-TARGETplus Non-Targeting Pool from Dharmacon. Transfected with Dharmafect1.

**Microarraying.** Performed on an Agilent platform, as described elsewhere[3].

**Immunoprecipitation, epitope mapping, western blotting.** Performed using standard protocols. Mouse M2 Anti-Flag antibody (Sigma), rabbit anti-human IgG–Fc horse radish peroxidase (HRP) conjugate (Pierce), and goat anti-rat IgG–HRP-conjugate (Pierce) were used. Protein G agarose beads (Millipore).



©2011 Macmillan Publishers Limited. All rights reserved

**Mass spectrometric analysis.** In all cases, samples were obtained by sequential immunoprecipitation, performed using anti-Flag affinity matrix (Sigma), 3×Flag peptide (Sigma) elution and re-precipitation using anti-haemagglutinin affinity matrix (Roche) as described[21]. For a detailed description of the mass spectrometric analysis see Supplementary Material.

**Surface plasmon resonance.** Performed by IBIS Technologies, Enschede, The Netherlands.

**Rat monoclonal antibodies anti-human LGR5.** Generated by Genovac by DNA vaccination with full-length human *LGR5* cDNA. Initial hybridoma screening was performed by fluorescence-activated cell sorting (FACS) on stable human LGR5 transfectants.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 29 March; accepted 27 June 2011.**
**Published online 4 July 2011.**

1. Barker, N. & Clevers, H. Leucine-rich repeat-containing G-protein-coupled receptors as markers of adult stem cells. *Gastroenterology* **138,** 1681–1696 (2010).
2. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* **449,** 1003–1007 (2007).
3. Barker, N. *et al.* Lgr5$^{+ve}$ stem cells drive self-renewal in the stomach and build long-lived gastric units *in vitro*. *Cell Stem Cell* **6,** 25–36 (2010).
4. Jaks, V. *et al.* Lgr5 marks cycling, yet long-lived, hair follicle stem cells. *Nature Genet.* **40,** 1291–1299 (2008).
5. Snippert, H. J. *et al. Lgr6* marks stem cells in the hair follicle that generate all cell lineages of the skin. *Science* **327,** 1385–1389 (2010).
6. Van Schoore, G., Mendive, F., Pochet, R. & Vassart, G. Expression pattern of the orphan receptor LGR4/GPR48 gene in the mouse. *Histochem. Cell Biol.* **124,** 35–50 (2005).
7. Mustata, R. C. *et al. Lgr4* is required for Paneth cell differentiation and maintenance of intestinal stem cells I. *EMBO Reports* **12,** 558–564 (2011).
8. van der Flier, L. G. *et al.* Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell* **136,** 903–912 (2009).
9. Morita, H. *et al.* Neonatal lethality of LGR5 null mice is associated with ankyloglossia and gastrointestinal distension. *Mol. Cell. Biol.* **24,** 9736–9743 (2004).
10. Kato, S. *et al.* Eye-open at birth phenotype with reduced keratinocyte motility in LGR4 null mice. *FEBS Lett.* **581,** 4685–4690 (2007).
11. Ireland, H., Houghton, C., Howard, L. & Winton, D. J. Cellular inheritance of a Cre-activated reporter gene to determine Paneth cell longevity in the murine small intestine. *Dev. Dyn.* **233,** 1332–1336 (2005).
12. Garcia, M. I. *et al.* LGR5 deficiency deregulates Wnt signaling and leads to precocious Paneth cell differentiation in the fetal intestine. *Dev. Biol.* **331,** 58–67 (2009).
13. van der Flier, L. G., Haegebarth, A., Stange, D. E., van de Wetering, M. & Clevers, H. OLFM4 is a robust marker for stem cells in human intestine and marks a subset of colorectal cancer cells. *Gastroenterology* **137,** 15–17 (2009).
14. Muncan, V. *et al.* Rapid loss of intestinal crypts upon conditional deletion of the Wnt/Tcf-4 target gene c-*Myc*. *Mol. Cell. Biol.* **26,** 8418–8426 (2006).
15. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche. *Nature* **459,** 262–265 (2009).
16. Pinto, D., Gregorieff, A., Begthel, H. & Clevers, H. Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes Dev.* **17,** 1709–1713 (2003).
17. Kuhnert, F. *et al.* Essential requirement for Wnt signaling in proliferation of adult small intestine and colon revealed by adenoviral expression of Dickkopf-1. *Proc. Natl Acad. Sci. USA* **101,** 266–271 (2004).
18. van Es, J. H. *et al.* Notch/γ-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. *Nature* **435,** 959–963 (2005).
19. Sansom, O. J. *et al. Myc* deletion rescues Apc deficiency in the small intestine. *Nature* **446,** 676–679 (2007).
20. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102,** 15545–15550 (2005).
21. Nakatani, Y. & Ogryzko, V. Immunoaffinity purification of mammalian protein complexes. *Methods Enzymol.* **370,** 430–444 (2003).
22. Kazanskaya, O. *et al.* R-Spondin2 is a secreted activator of Wnt/β-catenin signaling and is required for *Xenopus* myogenesis. *Dev. Cell* **7,** 525–534 (2004).
23. Kim, K. A. *et al.* R-spondin family members regulate the Wnt pathway by a common mechanism. *Mol. Biol. Cell* **19,** 2588–2596 (2008).
24. Kim, K. A. *et al.* Mitogenic influence of human R-spondin1 on the intestinal epithelium. *Science* **309,** 1256–1259 (2005).
25. Zhao, J. *et al.* R-spondin1, a novel intestinotrophic mitogen, ameliorates experimental colitis in mice. *Gastroenterology* **132,** 1331–1343 (2007).
26. Ohkawara, B., Glinka, A. & Niehrs, C. Rspo3 binds syndecan 4 and induces Wnt/PCP signaling via clathrin-mediated endocytosis to promote morphogenesis. *Dev. Cell* **20,** 303–314 (2011).
27. Nam, J. S., Turcotte, T. J., Smith, P. F., Choi, S. & Yoon, J. K. Mouse cristin/R-spondin family proteins are novel ligands for the Frizzled 8 and LRP6 receptors and activate β-catenin-dependent gene expression. *J. Biol. Chem.* **281,** 13247–13257 (2006).
28. Wei, Q. *et al.* R-spondin1 is a high affinity ligand for LRP6 and induces LRP6 phosphorylation and β-catenin signaling. *J. Biol. Chem.* **282,** 15903–15911 (2007).
29. Binnerts, M. E. *et al.* R-Spondin1 regulates Wnt signaling by inhibiting internalization of LRP6. *Proc. Natl Acad. Sci. USA* **104,** 14700–14705 (2007).
30. Glinka, A. *et al.* Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. *Nature* **391,** 357–362 (1998).
31. Korinek, V. *et al.* Constitutive transcriptional activation by a β-catenin-Tcf complex in APC$^{-/-}$ colon carcinoma. *Science* **275,** 1784–1787 (1997).
32. Flanagan, C. A. A. GPCR that is not "DRY". *Mol. Pharmacol.* **68,** 1–3 (2005).
33. Robine, S., Sahuquillo-Merino, C., Louvard, D. & Pringault, E. Regulatory sequences on the human villin gene trigger the expression of a reporter gene in a differentiating HT29 intestinal cell line. *J. Biol. Chem.* **268,** 11426–11434 (1993).
34. Klein, P. S. & Melton, D. A. A molecular mechanism for the effect of lithium on development. *Proc. Natl Acad. Sci. USA* **93,** 8455–8459 (1996).
35. Sato, T. *et al.* Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* **469,** 415–418 (2011).
36. Korinek, V. *et al.* Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nature Genet.* **19,** 379–383 (1998).
37. Ng, S. S. *et al.* Phosphatidylinositol 3-kinase signaling does not activate the Wnt cascade. *J. Biol. Chem.* **284,** 35308–35313 (2009).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank G. Vassart for *Lgr4-LacZ* intestinal tissue, D. Winton for *Ah-Cre* mice, S. Robine for *Villin-CreERT2* mice, A. Moerkamp and C. Verheul for experimental help and H. Farin for figures.

**Author Contributions** All Hubrecht Institute authors performed experiments under guidance of H.C.; S.M., A.J.R.H. and T.Y.L. performed mass spectrometry; P.K. and P.J.P. performed electron microscopy analysis; R.B.M.S. performed plasmon surface resonance; and Y.M. and K.N. generated the *Lgr4* knockout mouse.

**Author Information** Microarray data have been deposited in the GEO database under accession number GSE28265. Mass spectrometry data sets are available at ProteomeCommons.org Tranche Repository https://proteomecommons.org/tranche/data-downloader.jsp?h=2LOW5tCJBOfT%2FpcCAtMrPqCgTTOd247s6poPgSvwu16KiVwCfExWdJ0jifGdI4FraidTHUnI1PYhloT0nTs1zdwKmKEAAAAAAACzw%3D%3D. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to H.C. (h.clevers@hubrecht.eu).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Microarray analysis *Lgr4/5* knockout mice.** Small intestinal crypts were isolated 1 day after induction of deletion from *AHCre Lgr4fx/fx Lgr5fl/fl* as well as wild-type mice by incubation in 2 mM EDTA, RNA was isolated using TRIzol (Invitrogen) and 1 μg of RNA was labelled using a Quick Amp Labelling Kit, two colour (Agilent Technologies) with Cy5 and Cy3, respectively; all as described elsewhere[3]. Two separate biological replicates were performed in dye-swap, resulting in four individual arrays. Labelling, hybridization and washing were done according to Agilent guidelines. Differentially labelled cRNA was hybridized on 4X44K Agilent Whole Mouse Genome dual colour Microarrays (G4122F). Array data were normalized and retrieved using Feature Extraction (V.9.5.3, Agilent Technologies) and data analyses were performed using Microsoft Excel (Microsoft Corporation). Features were flagged, if signal intensities for both the Cy3 and Cy5 channel did not pass the feature extraction filter, 'significant and positive' or 'well above background'. Genes were considered downregulated if 4 out of 4 arrays showed a significant ($P < 0.05$) downregulation of $<-0.58$ (linear $-1.5$ fold). This resulted in 379 entries, which 307 unique genes (Supplementary Table array data will be available at the Gene Expression Omnibus http://www.ncbi.nlm.nih.gov/geo).

**R-spondin1 withdrawal in mouse intestinal organoids.** Crypts were isolated from a wild-type mouse small intestine by incubating with 2 mM EDTA in PBS for 30 min at 4 °C, and subsequently grown in crypt culture medium as reported previously[16]. Briefly, isolated crypts were cultured in Matrigel (BD Bioscience) in 24-well plates, and advanced DMEM/F12 medium (Invitrogen) containing epithelial growth factor, noggin and Rspo1 was added after polymerization of Matrigel. Confluent organoids were split into multiple wells, and were then cultured in crypt culture medium in the presence or absence of Rspo1. One day after Rspo1 withdrawal organoids (−RSpo1 organoids) and the control organoids (+Rspo1 organoids) were then collected for RNA extraction and microarray analysis. RNA from control and Rspo1 depleted organoids (1 μg), together with universal mouse reference RNA (Stratagene), was labelled using a Quick Amp Labelling Kit, two colour (Agilent Technologies) with Cy5 and Cy3, respectively. Samples were hybridized to 4X44K Whole Mouse Genome Microarrays (Agilent, G4122F) according to manufacturer's instructions. Microarray signal and background information were retrieved and normalized using the Feature Extraction program (V.9.5.3, Agilent Technologies). Samples were considered as well-measured when the fluorescent signals in red channel (Cy5) in either of the samples were greater than twofold above the local background. Differences between −Rspo1 organoids and +Rspo1 organoids were calculated by subtracting the ratio for '−Rspo1 organoids versus reference RNA' from '+Rspo1 versus reference RNA'. Array data is available at the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo).

**Heat map and gene set enrichment analysis.** Heat maps were generated using MeV (Multiple Experiment Viewer v.4.3)[42]. Heat maps were generated using the 306 genes from the *Lgr4/5* gene set and plotting the ratios of two different experiments for these genes. Figure 2a contains the ratios from ref. 15, where the authors deleted *Apc* in the *Ah-Cre* $Apc^{fl/fl}$ mice and performed microarray analysis 3 days after deletion. Figure 2b contains the ratios form the Rspo1 withdrawal experiment described above. Gene set enrichment analysis (GSEA) implemented with GSEAP v.2.0 (http://www.broad.mit.edu/gsea) was used to identify significant enrichments of the *Lgr4/5* gene set in the two different experimental scenarios mentioned above. All 'well measured' features ($n = 20{,}844$) for the Rspo1 experiment from the Agilent arrays and all features with a ratio ($n = 45{,}101$) for the *Apc* arrays from the Affymetrix platform were used for the ranked gene list in GSEA.

**Mass spectrometric analysis.** The samples were subjected to SDS–PAGE followed by Coomassie blue staining. The gel lane was sliced into 24 equal sections and digested. In brief, protein reduction and alkylation was performed with dithiothreitol (60 °C, 1 h) and iodoacetamide (dark, room temperature, 30 min), respectively. Digestion was performed with trypsin overnight at 37 °C. Peptides were extracted with 10% formic acid[38].

The extracted peptides were subjected to nanoscale liquid chromatography tandem mass spectrometry (nanoLC-MS/MS) analysis, performed on an Agilent 1200 HPLC system (Agilent technologies) connected to an LTQ Orbitrap Velos (ThermoFisher)[39]. The nanoLC was equipped with a 20 mm × 100 μm internal diameter trap column and a 400 mm × 50 μm internal diameter analytical column (Reprosil C18, Dr Maisch). Trapping was performed at a flow of 5 μl $min^{-1}$ for 10 min and the fractions were eluted using a 45 min linear gradient from 0 to 40% solvent B (0.1 M acetic acid in 80% acetonitrile (v/v), in which solvent A was 0.1 M acetic acid), 40 to 100% solvent B in 2 min and 100% B for 2.5 min. The analytical flow rate was 100 nl $min^{-1}$ and the column effluent was directly introduced into the electrospray source of the mass spectrometer using a standard coated fused silica emitter (New Objective) (outer diameter 360 μm, tip internal diameter 10 μm) biased to 1.7 kV. The mass spectrometer was operated in positive ion mode and in data-dependent mode to automatically switch between MS and MS/MS. The five most intense ions in the survey scan were fragmented in the linear ion trap using collision-induced dissociation[40]. The target ion setting was $5 \times 10^5$ for the Orbitrap, with a maximum fill-time of 250 ms. Fragment ion spectra were acquired in the LTQ with an automatic gain control value of $5 \times 10^3$ and a max injection time of 100 ms.

Protein identification: raw mass spectrometry data were converted to peak lists using MaxQuant version 1.0.13.13 (ref. 41). Spectra were searched against the IPI (International Protein Index) Human database version 3.37 (69,164 sequences; 29,064,824 residues) using the Mascot search engine (version 2.3.02; http://www.matrixscience.com), with trypsin set as enzyme. The database search was made with the following parameters set to consider a peptide tolerance of ±15 p.p.m., a fragment tolerance of ±0.5 Da, allowing two missed cleavages, carbamidomethyl (C) as fixed modification; oxidation (M) and protein N-terminal acetylation as variable modifications.

**Plasmon surface resonance.** A pre-activated ester sensor chip (IBIS Technologies) was spotted using a Continuous Flow Microspotter (Wasatch Microfluidics). In total, 32 spots were created with both mouse anti-Flag and goat anti-human IgG in 8× serial dilution (start concentrations were 500 μg $ml^{-1}$ and 100 μg $ml^{-1}$, respectively) in 10 mM MES buffer pH 5.5. Control (reference) spots contained human serum albumin, anti-human serum albumin and MES buffer. After preparing the sensor chip, the slider was positioned in an IBIS MX96 (IBIS Technologies) for label-free surface plasmon resonance (SPR) array analysis. The instrument enables multiplex interactions up to 96-plex using scanning imaging optics for automated calculation of the SPR dip shifts of all region of interests simultaneously. The signal-to-noise ratio of the instrument which reflects the limit of detection is better than 1 RU corresponding to 1 picogram protein per square millimetre. In the IBIS MX96, back and forth mixing is applied enabling minimal use of sample while the length of the exposure of the sample to the microarray is unlimited and not affected by the volume of, for example, an injection loop. A two-step interaction process was carried out and the multiplex interaction event to all spots of the array was recorded simultaneously. In this way, chip-to-chip and sample-to-sample variations can be excluded, while positive and negative controls and referencing can be applied instantly. A microscope image of the sensor chip (not shown here) can reveal any irregularities, heterogeneities of the spots and/or disturbing air bubbles. First 70 μl of RSPO1–FH was injected followed by dissociation and second injection of 70 μl LGR5–Fc. After regeneration with acid buffer (Gly-HCl, 10 mM, pH 2.0) for 120 s, the control experiment was capture of RSPO1–FH followed by injection of noggin–Fc (results see Fig. 4b). For affinity constant determination first RSPO1–FH was captured on the anti-Flag spot until saturation of the spot followed by injection of LGR5–Fc (8 μg $ml^{-1}$). RSPO1 was not immobilized directly on the chip because it did not survive the acid regeneration process (data not shown). Capturing of RSPO1–FH was possible on the anti-Flag antibody spot from growth medium. Supplementary Fig. 1 is an overlay plot of three interaction series of first RSPO1–FH (2 min association was sufficient to reach saturation) and three different concentrations of LGR5–Fc (molecular mass 176 kDa) corresponding to 45 nM, 23 nM and 11 nM injections. Referencing was carried out by subtraction of the anti-Flag spot signal with the signal coming from a HAS-loaded spot in SPRint software (IBIS Technologies) for compensating bulk refractive index shifts for example, as a result of growth medium compounds. The affinity constant was calculated using a discrete 1:1 interaction model using global fitting (Scrubber 2, BioLogic Software Pty Ltd). In Supplementary Fig. 1, the residual plot of the experimental curve minus fit values is shown and although it cannot be revealed that the interaction process is according to a discrete 1:1 interaction model the affinity constant was calculated to be ~3 nM. It was not necessary to add an extra fit parameter for mass transport limitation compensation, due to a high back and forth mixing condition in the IBIS MX96 flow cell.


38. Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68,** 850–858 (1996).
39. Raijmakers, R. *et al.* Automated online sequential isotope labeling for protein quantitation applied to proteasome tissue-specific diversity. *Mol. Cell. Proteomics* **7,** 1755–1762 (2008).
40. Frese, C. K. *et al.* Improved peptide identification by targeted fragmentation using CID, HCD and ETD on a LTQ-Orbitrap Velos. *J. Proteome Res.* **10,** 2377–2388 (2011).
41. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnol.* **26,** 1367–1372 (2008).
42. Saeed, A. I. *et al.* TM4: a free, open-source system for microarray data management and analysis. Biotechniques, **34,** 374–378 (2003).


©2011 Macmillan Publishers Limited. All rights reserved

ARTICLE

doi:10.1038/nature10351

# Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma

Ryan D. Morin[1]*, Maria Mendez-Lago[1]*, Andrew J. Mungall[1], Rodrigo Goya[1], Karen L. Mungall[1], Richard D. Corbett[1], Nathalie A. Johnson[2], Tesa M. Severson[1], Readman Chiu[1], Matthew Field[1], Shaun Jackman[1], Martin Krzywinski[1], David W. Scott[2], Diane L. Trinh[1], Jessica Tamura-Wells[1], Sa Li[1], Marlo R. Firme[1], Sanja Rogic[2], Malachi Griffith[1], Susanna Chan[1], Oleksandr Yakovenko[1], Irmtraud M. Meyer[3], Eric Y. Zhao[1], Duane Smailus[1], Michelle Moksa[1], Suganthi Chittaranjan[1], Lisa Rimsza[4], Angela Brooks-Wilson[1,5], John J. Spinelli[6,7], Susana Ben-Neriah[2], Barbara Meissner[2], Bruce Woolcock[2], Merrill Boyle[2], Helen McDonald[1], Angela Tam[1], Yongjun Zhao[1], Allen Delaney[1], Thomas Zeng[1], Kane Tse[1], Yaron Butterfield[1], Inanç Birol[1], Rob Holt[1], Jacqueline Schein[1], Douglas E. Horsman[2], Richard Moore[1], Steven J. M. Jones[1], Joseph M. Connors[2], Martin Hirst[1], Randy D. Gascoyne[2,8] & Marco A. Marra[1,9]

**Follicular lymphoma (FL) and diffuse large B-cell lymphoma (DLBCL) are the two most common non-Hodgkin lymphomas (NHLs). Here we sequenced tumour and matched normal DNA from 13 DLBCL cases and one FL case to identify genes with mutations in B-cell NHL. We analysed RNA-seq data from these and another 113 NHLs to identify genes with candidate mutations, and then re-sequenced tumour and matched normal DNA from these cases to confirm 109 genes with multiple somatic mutations. Genes with roles in histone modification were frequent targets of somatic mutation. For example, 32% of DLBCL and 89% of FL cases had somatic mutations in *MLL2*, which encodes a histone methyltransferase, and 11.4% and 13.4% of DLBCL and FL cases, respectively, had mutations in *MEF2B*, a calcium-regulated gene that cooperates with CREBBP and EP300 in acetylating histones. Our analysis suggests a previously unappreciated disruption of chromatin biology in lymphomagenesis.**

Non-Hodgkin lymphomas (NHLs) are cancers of B, T or natural killer lymphocytes. The two most common types of NHL, follicular lymphoma (FL) and diffuse large B-cell lymphoma (DLBCL), together comprise 60% of new B-cell NHL diagnoses each year in North America[1]. FL is an indolent and typically incurable disease characterized by clinical and genetic heterogeneity. DLBCL is aggressive and likewise heterogeneous, comprising at least two distinct subtypes that respond differently to standard treatments. Both FL and the germinal centre B-cell (GCB) cell of origin (COO) subtype of DLBCL derive from germinal centre B cells, whereas the activated B-cell (ABC) variety, which has a more aggressive clinical course, is thought to originate from B cells that have exited, or are poised to exit, the germinal centre[2]. Current knowledge of the specific genetic events leading to DLBCL and FL is limited to the presence of a few recurrent genetic abnormalities[2]. For example, 85–90% of FL and 30–40% of GCB DLBCL cases[3,4] harbour t(14;18)(q32;q21), which results in deregulated expression of the BCL2 oncoprotein. Other genetic abnormalities unique to GCB DLBCL include amplification of the *c-REL* gene and of the miR-17-92 microRNA cluster[5]. In contrast to GCB cases, 24% of ABC DLBCLs harbour structural alterations or inactivating mutations affecting *PRDM1*, which is involved in differentiation of GCB cells into antibody-secreting plasma cells[6]. ABC-specific mutations also affect genes regulating NF-κB signalling[7,8,9], with *TNFAIP3* (also known as *A20*) and *MYD88* (ref. 10) the most abundantly mutated in 24% and 39% of cases, respectively. To enhance our understanding of the genetic architecture of B-cell NHL, we undertook a study to (1) identify somatic mutations and (2) determine the prevalence, expression and focal recurrence of mutations in FL and DLBCL. Using strategies and techniques applied to cancer genome and transcriptome characterization by ourselves and others[11,12,13], we sequenced tumour DNA and/or RNA from 117 tumour samples and 10 cell lines (Supplementary Tables 1 and 2) and identified 651 genes (Supplementary Figure 1) with evidence of somatic mutation in B-cell NHL. After validation, we showed that 109 genes were somatically mutated in two or more NHL cases. We further characterized the frequency and nature of mutations within *MLL2* and *MEF2B*, which were among the most frequently mutated genes with no previously known role in lymphoma.

## Identification of recurrently mutated genes

We sequenced the genomes or exomes of 14 NHL cases, all with matched constitutional DNA sequenced to comparable depths (Supplementary Tables 1 and 2). After screening for single nucleotide variants followed by subtraction of known polymorphisms and visual inspection of the sequence read alignments, we identified 717 non-synonymous variants (coding single nucleotide variants; cSNVs) affecting 651 genes (Supplementary Figure 1 and Supplementary Methods). We identified between 20 and 135 cSNVs in each of these genomes. Only 25 of the 651 genes with cSNVs were represented in the cancer gene census (December 2010 release)[14].

We performed RNA sequencing (RNA-seq) on these 14 NHL cases and an expanded set of 113 samples comprising 83 DLBCL, 12 FL and 8 B-cell NHL cases with other histologies and 10 DLBCL-derived cell lines (Supplementary Table 2). We analysed these data to identify

[1]Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia V5Z 1L3, Canada. [2]Centre for Lymphoid Cancer, BC Cancer Agency, Vancouver, British Columbia V5Z 1L3, Canada. [3]Centre for High-throughput Biology, Department of Computer Science, Vancouver, British Columbia V6T 1Z4, Canada. [4]Department of Pathology, University of Arizona, Tucson, Arizona 85724, USA. [5]Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada. [6]Cancer Control Research, BC Cancer Agency, Vancouver, British Columbia V5Z 1L3, Canada. [7]School of Population and Public Health, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada. [8]Department of Pathology, University of British Columbia, Vancouver, British Columbia V6T 2B5, Canada. [9]Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia V6H 3N1, Canada.
*These authors contributed equally to this work.



©2011 Macmillan Publishers Limited. All rights reserved

novel fusion transcripts (Supplementary Table 3) and cSNVs (Fig. 1). We identified 240 genes with at least one cSNV in a genome/exome or an RNA-seq 'mutation hot spot' (see later), and with cSNVs in at least three cases in total (Supplementary Table 4). We selected cSNVs from each of these 240 genes for re-sequencing to confirm their somatic status. We did not re-sequence genes with previously documented mutations in lymphoma (for example, *CD79B*, *BCL2*). We confirmed the somatic status of 543 cSNVs in 317 genes, with 109 genes having at least two confirmed somatic mutations (Supplementary Tables 4 and 5). Of the successfully re-sequenced cSNVs predicted from the genomes, 171 (94.5%) were confirmed somatic, 7 were false calls and 3 were present in the germ line. These 109 recurrently mutated genes were significantly enriched for genes implicated in lymphocyte activation ($P = 8.3 \times 10^{-4}$; for example, *STAT6*, *BCL10*), lymphocyte differentiation ($P = 3.5 \times 10^{-3}$; for example, *CARD11*), and regulation of apoptosis ($P = 1.9 \times 10^{-3}$; for example, *BTG1*, *BTG2*). Also significantly enriched were genes linked to transcriptional regulation ($P = 5.4 \times 10^{-4}$; for example, *TP53*) and genes involved in methylation ($P = 2.2 \times 10^{-4}$) and acetylation ($P = 1.2 \times 10^{-2}$), including histone methyltransferase (HMT) and acetyltransferase (HAT) enzymes known previously to be mutated in lymphoma (for example, *EZH2* (ref. 13) and *CREBBP* (ref. 15); Supplementary Methods).

Mutation hot spots can result from mutations at sites under strong selective pressure and we have previously identified such sites using RNA-seq data[13]. We searched our RNA-seq data for genes with mutation hot spots, and identified 10 genes that were not mutated in the 14 genomes (*PIM1*, *FOXO1*, *CCND3*, *TP53*, *IRF4*, *BTG2*, *CD79B*, *BCL7A*, *IKZF3* and *B2M*), of which five (*FOXO1*, *CCND3*, *BTG2*, *IKZF3* and *B2M*) were not previously known targets of point mutation in NHL (Supplementary Table 6 and Supplementary Methods). *FOXO1*, *BCL7A* and *B2M* had hot spots affecting their start codons. The effect of a *FOXO1* start codon mutation, which was observed in three cases, was further studied using a cell line in which the initiating ATG was mutated to TTG. Western blots probed with a FOXO1 antibody revealed a band with a reduced molecular weight, indicative of a FOXO1 amino-terminal truncation (Supplementary Figure 2), consistent with use of the next in-frame ATG for translation initiation. A second hot spot in *FOXO1* at T24 was mutated in two cases. T24 is reportedly phosphorylated by AKT subsequent to B-cell receptor (BCR) stimulation[16] inducing FOXO1 nuclear export.

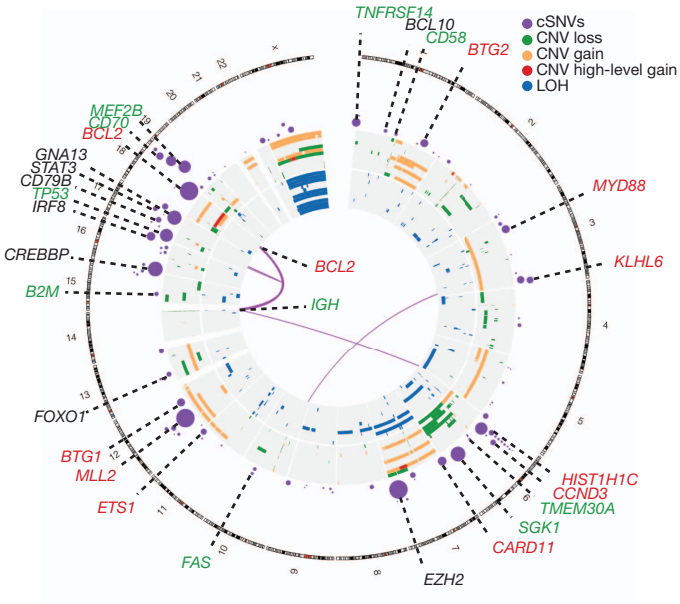



**Figure 1 | Genome-wide visualization of somatic mutation targets in NHL.** Overview of structural rearrangements and copy number variations (CNVs) in the 11 DLBCL genomes and cSNVs in the 109 recurrently mutated genes identified in our analysis. Inner arcs represent somatic fusion transcripts identified in at least one of the 11 genomes. The CNVs and LOH detected in each of the 11 DLBCL tumour/normal pairs are displayed on the concentric sets of rings. The inner 11 rings show regions of enhanced homozygosity plotted with blue (interpreted as LOH). The outer 11 rings show somatic CNVs. Purple circles indicate the position of genes with at least two confirmed somatic mutations with circle diameter proportional to the number of cases with cSNVs detected in that gene. Circles representing the genes with significant evidence for positive selection are labelled. Coincidence between recurrently mutated genes and regions of gain/loss are colour-coded in the labels (green, loss; red, gain). For example *B2M*, which encodes beta-2-microglobulin, is recurrently mutated and is deleted in two cases.

We analysed the RNA-seq data to determine whether any of the somatic mutations in the 109 recurrently mutated genes showed evidence for allelic imbalance with expression favouring one allele. Out of 380 expressed heterozygous mutant alleles, we observed preferential expression of the mutation for 16.8% (64/380) and preferential expression of the wild type for 27.8% (106/380; Supplementary Table 7). Seven genes showed evidence for significant preferential expression of the mutant allele in at least two cases: *BCL2*, *CARD11*, *CD79B*, *EZH2*, *IRF4*, *MEF2B* and *TP53*; Supplementary Methods. In 27 out of 43 cases with *BCL2* cSNVs, expression favoured the mutant allele, consistent with the previously-described hypothesis that the translocated (and hence, transcriptionally deregulated) allele of *BCL2* is targeted by somatic hypermutation[17]. Examples of mutations at known oncogenic hot spot sites such as F123I in *CARD11* (ref. 18) showed allelic imbalance favouring the mutant allele in some cases. Similarly, we noted expression favouring two novel hot spot mutations in *MEF2B* (Y69 and D83) and two sites in *EZH2* not previously reported as mutated in lymphoma (A682G and A692V).

We sought to distinguish new cancer-related mutations from passenger mutations using the approach proposed previously[19]. We reasoned that this would reveal genes with strong selection signatures, and mutations in such genes would be good candidate cancer drivers. We identified 26 genes with significant evidence for positive selection (false discovery rate = 0.03, Supplementary Methods), with either selective pressure for acquiring non-synonymous point mutations or truncating/nonsense mutations (Supplementary Methods; Table 1 and Supplementary Table 8). Included were known lymphoma oncogenes (*BCL2*, *CD79B* (ref. 9), *CARD11* (ref. 18), *MYD88* (ref. 10) and *EZH2* (ref. 13)), all of which showed signatures indicative of selection for non-synonymous variants.

## Evidence for selection of inactivating changes

We expected tumour suppressor genes to show strong selection for the acquisition of nonsense mutations. In our analysis, the eight most significant genes included seven with strong selective pressure for nonsense mutations, including the known tumour suppressor genes *TP53* and *TNFRSF14* (ref. 20 ; Table 1). *CREBBP*, recently reported as commonly inactivated in DLBCL[15], also showed some evidence for acquisition of nonsense mutations and cSNVs (Supplementary Figure 3 and Supplementary Table 9). We also observed enrichment for nonsense mutations in *BCL10*, a positive regulator of NF-κB, in which oncogenic truncated products have been described in lymphomas[21]. The remaining strongly significant genes (*BTG1*, *GNA13*, *SGK1* and *MLL2*) had no reported role in lymphoma. *GNA13* was affected by mutations in 22 cases including multiple nonsense mutations. *GNA13* encodes the alpha subunit of a heterotrimeric G-protein coupled receptor responsible for modulating RhoA activity[22]. Some of the mutated residues negatively affect its function[23,24], including a T203A mutation, which also showed allelic imbalance favouring the mutant allele (Supplementary Table 7). GNA13 protein was reduced or absent on western blots in cell lines harbouring either a nonsense mutation, a stop codon deletion, a frame shifting deletion, or changes affecting splice sites (Supplementary Methods and Supplementary Figure 4).



©2011 Macmillan Publishers Limited. All rights reserved

**Table 1 | Overview of cSNVs and confirmed somatic mutations in most frequently mutated genes**

| Gene | Cases | | | Total | | | Somatic cSNVs (RNA-seq cohort)* | *P* (raw) | *q* | NS SP | T SP | Skew (M, WT, both)† |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | NS | S | T | NS | S | T | | | | | | |
| *MLL2*‡ | 16 | 8 | 17 | 17 | 8 | 18 | 10 | $\mathbf{6.85 \times 10^{-8}}$ | $8.50 \times 10^{-7}$ | 0.834 | 14.4 | WT |
| *TNFRSF14* [G]‡ | 7 | 1 | 7 | 8 | 1 | 7 | 11 | $\mathbf{6.85 \times 10^{-8}}$ | $8.50 \times 10^{-7}$ | 7.52 | 118 | Both |
| *SGK1* [G]‡ | 18 | 6 | 6 | 37 | 10 | 6 | 9 | $\mathbf{6.85 \times 10^{-8}}$ | $8.50 \times 10^{-7}$ | 19.5 | 61.7 | – |
| *BCL10*‡ | 2 | 0 | 4 | 3 | 0 | 4 | 4 | $\mathbf{6.85 \times 10^{-8}}$ | $8.50 \times 10^{-7}$ | 3.62 | 112 | WT |
| *GNA13* [G]‡ | 21 | 1 | 2 | 33 | 1 | 2 | 5 | $\mathbf{6.85 \times 10^{-8}}$ | $8.50 \times 10^{-7}$ | 24.1 | 25.7 | Both |
| *TP53* [G]‡ | 20 | 2 | 1 | 23 | 3 | 1 | 22 | $\mathbf{6.85 \times 10^{-8}}$ | $8.50 \times 10^{-7}$ | 15.6 | 14.1 | Both |
| *EZH2* [G]‡ | 33 | 0 | 0 | 33 | 0 | 0 | 33 | $\mathbf{6.85 \times 10^{-8}}$ | $8.50 \times 10^{-7}$ | 11.4 | 0.00 | Both |
| *BTG2*‡ | 12 | 6 | 1 | 14 | 6 | 1 | 2 | $6.85 \times 10^{-8}$ | $8.50 \times 10^{-7}$ | 23.9 | 35.1 | – |
| *BCL2* [G]‡ | 42 | 45 | 0 | 96 | 105 | 0 | 43 | $\mathbf{9.35 \times 10^{-8}}$ | $8.50 \times 10^{-7}$ | 3.78 | 0.00 | M |
| *BCL6*‡§ | 11 | 2 | 0 | 12 | 2 | 0 | 2 | $\mathbf{9.35 \times 10^{-8}}$ | $8.50 \times 10^{-7}$ | 0.175 | 0.00 | M |
| *CIITA*‡§ | 5 | 3 | 0 | 6 | 3 | 0 | 2 | $\mathbf{9.35 \times 10^{-8}}$ | $8.50 \times 10^{-7}$ | 0.086 | 0.00 | |
| *FAS*‡ | 2 | 0 | 4 | 3 | 0 | 4 | 2 | $1.52 \times 10^{-7}$ | $1.17 \times 10^{-6}$ | 2.54 | 66.5 | WT |
| *BTG1*‡ | 11 | 6 | 2 | 11 | 7 | 2 | 10 | $1.52 \times 10^{-7}$ | $1.17 \times 10^{-6}$ | 17.5 | 52.5 | Both |
| *MEF2B* [G]‡ | 20 | 2 | 0 | 20 | 2 | 0 | 10 | $2.05 \times 10^{-7}$ | $1.47 \times 10^{-6}$ | 14.2 | 0.00 | M |
| *IRF8*‡ | 11 | 5 | 3 | 14 | 5 | 3 | 3 | $4.55 \times 10^{-7}$ | $3.03 \times 10^{-6}$ | 8.82 | 28.2 | WT |
| *TMEM30A*‡ | 1 | 0 | 4 | 1 | 0 | 4 | 4 | $6.06 \times 10^{-7}$ | $3.79 \times 10^{-6}$ | 0.785 | 65.0 | WT |
| *CD58*‡ | 2 | 0 | 3 | 2 | 0 | 3 | 2 | $2.42 \times 10^{-6}$ | $1.43 \times 10^{-5}$ | 2.29 | 69.2 | – |
| *KLHL6*‡ | 10 | 2 | 2 | 12 | 2 | 2 | 4 | $1.00 \times 10^{-5}$ | $5.26 \times 10^{-5}$ | 5.42 | 16.4 | – |
| *MYD88* [A]‡ | 13 | 2 | 0 | 14 | 2 | 0 | 9 | $1.00 \times 10^{-5}$ | $5.26 \times 10^{-5}$ | 12.4 | 0.00 | WT |
| *CD70*‡ | 5 | 0 | 1 | 5 | 0 | 2 | 3 | $1.70 \times 10^{-5}$ | $8.48 \times 10^{-5}$ | 7.08 | 44.0 | – |
| *CD79B* [A]‡ | 7 | 2 | 1 | 9 | 2 | 1 | 5 | $2.00 \times 10^{-5}$ | $9.52 \times 10^{-5}$ | 10.9 | 18.3 | M |
| *CCND3*‡ | 7 | 1 | 2 | 7 | 1 | 2 | 6 | $2.80 \times 10^{-5}$ | $1.27 \times 10^{-4}$ | 6.55 | 36.3 | WT |
| *CREBBP*‡ | 20 | 7 | 4 | 24 | 7 | 4 | 9 | $1.00 \times 10^{-4}$ | $4.35 \times 10^{-4}$ | 2.72 | 6.04 | Both |
| *HIST1H1C*‡ | 9 | 0 | 0 | 10 | 0 | 0 | 6 | $1.80 \times 10^{-4}$ | $7.50 \times 10^{-4}$ | 11.9 | 0.00 | Both |
| *B2M*‡ | 7 | 0 | 0 | 7 | 0 | 0 | 4 | $3.90 \times 10^{-4}$ | $1.56 \times 10^{-3}$ | 16.6 | 0.00 | WT |
| *ETS1*‡ | 10 | 1 | 0 | 10 | 1 | 0 | 4 | $4.10 \times 10^{-4}$ | $1.58 \times 10^{-3}$ | 5.76 | 0.00 | WT |
| *CARD11*‡ | 14 | 3 | 0 | 14 | 3 | 0 | 3 | $1.90 \times 10^{-3}$ | $7.04 \times 10^{-3}$ | 3.37 | 0.00 | Both |
| *FAT2*‡§ | 2 | 1 | 0 | 2 | 1 | 0 | 2 | $6.30 \times 10^{-3}$ | $2.25 \times 10^{-2}$ | 0.128 | 0.00 | – |
| *IRF4*‡§ | 9 | 4 | 0 | 26 | 5 | 0 | 5 | $7.00 \times 10^{-3}$ | $2.41 \times 10^{-2}$ | 0.569 | 0.00 | Both |
| *FOXO1*‡ | 8 | 4 | 0 | 10 | 4 | 0 | 4 | $7.60 \times 10^{3}$ | $2.53 \times 10^{-2}$ | 4.02 | 0.00 | – |
| *STAT3* | 9 | 0 | 0 | 9 | 0 | 0 | 4 | $2.19 \times 10^{-2}$ | $6.08 \times 10^{-2}$ | – | – | Both |
| *RAPGEF1* | 8 | 3 | 0 | 10 | 3 | 0 | 3 | $2.98 \times 10^{-2}$ | $7.45 \times 10^{-2}$ | – | – | WT |
| *ABCA7* | 12 | 3 | 0 | 15 | 3 | 0 | 2 | $7.76 \times 10^{-2}$ | $1.67 \times 10^{-1}$ | – | – | WT |
| *RNF213* | 10 | 8 | 0 | 10 | 8 | 0 | 2 | $7.87 \times 10^{-2}$ | $1.67 \times 10^{-1}$ | – | – | – |
| *MUC16* | 17 | 12 | 0 | 39 | 25 | 0 | 2 | $8.32 \times 10^{-2}$ | $1.73 \times 10^{-1}$ | – | – | – |
| *HDAC7* | 8 | 4 | 0 | 8 | 4 | 0 | 2 | $8.94 \times 10^{-2}$ | $1.82 \times 10^{-1}$ | – | – | WT |
| *PRKDC* | 7 | 3 | 0 | 7 | 4 | 0 | 2 | $1.06 \times 10^{-1}$ | $2.05 \times 10^{-1}$ | – | – | – |
| *SAMD9* | 9 | 2 | 0 | 9 | 2 | 0 | 2 | $1.79 \times 10^{-1}$ | $3.01 \times 10^{-1}$ | – | – | – |
| *TAF1* | 10 | 0 | 0 | 10 | 0 | 0 | 2 | $3.03 \times 10^{-1}$ | $4.74 \times 10^{-1}$ | – | – | – |
| *PIM1* | 20 | 19 | 0 | 33 | 34 | 0 | 11 | $3.40 \times 10^{-1}$ | $5.23 \times 10^{-1}$ | – | – | WT |
| *COL4A2* | 8 | 2 | 0 | 8 | 2 | 0 | 2 | $7.64 \times 10^{-1}$ | $8.99 \times 10^{-1}$ | – | – | – |
| *EP300* | 8 | 7 | 1 | 8 | 7 | 1 | 3 | $9.54 \times 10^{-1}$ | 1.00 | – | – | WT |

Individual cases with non-synonymous (NS), synonymous (S) and truncating (T) mutations and the total number of mutations of each class are shown separately because some genes contained multiple mutations in the same case. The *P* values indicated in bold are the upper limit on the *P* value for that gene determined with the approach described in ref. 19 (see Supplementary Methods), *q* is the Benjamini-corrected *q* value, and NS SP and T SP refer to selective pressure estimates from this model for the acquisition of non-synonymous or truncating mutations, respectively. Genes with a superscript of either A or G were found to have mutations significantly enriched in ABC or GCB cases, respectively ($P < 0.05$, Fisher's exact test).

* Additional somatic mutations identified in larger cohorts and insertion/deletion mutations are not included in this total.

† 'Both' indicates that we observed separate cases in which skewed expression was seen but where this skew was not consistent for the mutant or wild-type allele.

‡ Genes significant at a false discovery rate of 0.03. SNVs in BCL2 and previously confirmed hot spot mutations in *EZH2* and *CD79B* are probably somatic in these samples based on published observations of others.

§ Selective pressure estimates are both < 1 indicating purifying selection rather than positive selection acting on this gene.

*SGK1* encodes a phosphatidylinositol-3-OH kinase (PI(3)K)-regulated kinase with functions including regulation of FOXO transcription factors[25], regulation of NF-κB by phosphorylating IκB kinase[26], and negative regulation of NOTCH signalling[27]. *SGK1* also resides within a region of chromosome 6 commonly deleted in DLBCL (Fig. 1)[5]. The mechanism by which *SGK1* and *GNA13* inactivation may contribute to lymphoma is unclear, but the strong degree of apparent selection towards their inactivation and their overall high mutation frequency (each mutated in 18 of 106 DLBCL cases) suggests that their loss contributes to B-cell NHL. Certain genes are known to be mutated more commonly in GCB DLBCLs (for example, *TP53* (ref. 28) and *EZH2* (ref. 13)). Here, both *SGK1* and *GNA13* mutations were found only in GCB cases ($P = 1.93 \times 10^{-3}$ and $2.28 \times 10^{-4}$, Fisher's exact test; $n = 15$ and 18, respectively) (Fig. 2). Two additional genes (*MEF2B* and *TNFRSF14*) with no previously described role in DLBCL showed a similar restriction to GCB cases (Fig. 2).

## Inactivating *MLL2* mutations

*MLL2* showed the most significant evidence for selection and the largest number of nonsense SNVs. Our RNA-seq analysis indicated that 26.0% (33/127) of cases carried at least one *MLL2* cSNV. To address the possibility that variable RNA-seq coverage of *MLL2* failed to capture some mutations, we PCR-amplified the entire *MLL2* locus (~36 kilobases) in 89 cases (35 primary FLs, 17 DLBCL cell lines, and 37 DLBCLs). Of these cases 58 were among the RNA-seq cohort. Illumina amplicon re-sequencing (Supplementary Methods) revealed 78 mutations, confirming the RNA-seq mutations in the overlapping cases and identifying 33 additional mutations. We confirmed the somatic status of 46 variants using Sanger sequencing (Supplementary Table 10), and showed that 20 of the 33 additional mutations were insertions or deletions (indels). Three SNVs at splice sites were also detected, as were 10 new cSNVs that had not been detected by RNA-seq.

The somatic mutations were distributed across *MLL2* (Fig. 3a). Of these, 37% ($n = 29/78$) were nonsense mutations, 46% ($n = 36/78$) were indels that altered the reading frame, 8% ($n = 6/78$) were point mutations at splice sites and 9% ($n = 7/78$) were non-synonymous amino acid substitutions (Table 2). Four of the somatic splice site mutations had effects on *MLL2* transcript length and structure. For example, two heterozygous splice site mutations resulted in the use of a novel splice donor site and an intron retention event.

Approximately half of the NHL cases we sequenced had two *MLL2* mutations (Supplementary Table 10). We used bacterial artificial



©2011 Macmillan Publishers Limited. All rights reserved

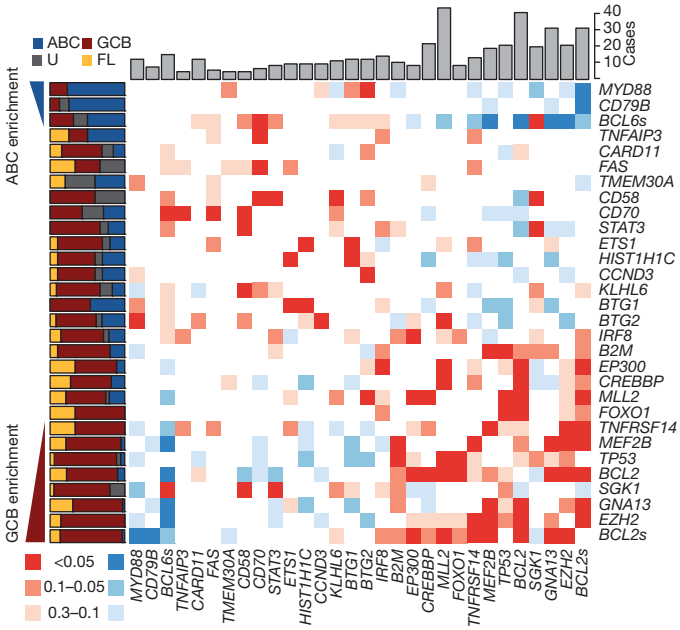



**Figure 2 | Overview of mutations and potential cooperative interactions in NHL.** This heat map displays possible trends towards co-occurrence (red) and mutual exclusion (blue) of somatic mutations and structural rearrangements. Colours were assigned by taking the minimum value of a left- and right-tailed Fisher's exact test. To capture trends a *P*-value threshold of 0.3 was used, with the darkest shade of the colour indicating those meeting statistical significance ($P \leq 0.05$). The relative frequency of mutations in ABC (blue), GCB (red), unclassifiable (black) DLBCLs and FL (yellow) cases is shown on the left. Genes were arranged with those having significant ($P < 0.05$, Fisher's exact test) enrichment for mutations in ABC cases (blue triangle) towards the top (and left) and those with significant enrichment for mutations in GCB cases (red triangle) towards the bottom (and right). The total number of cases in which each gene contained either cSNVs or confirmed somatic mutations is shown at the top. The cluster of blue squares (upper-right) results from the mutual exclusion of the ABC-enriched mutations (for example, *MYD88*, *CD79B*) from the GCB-enriched mutations (for example, *EZH2*, *GNA13*). Presence of structural rearrangements involving the two oncogenes *BCL6* and *BCL2* (indicated as *BCL6*s and *BCL2*s) was determined with FISH techniques using break-apart probes (Supplementary Methods).

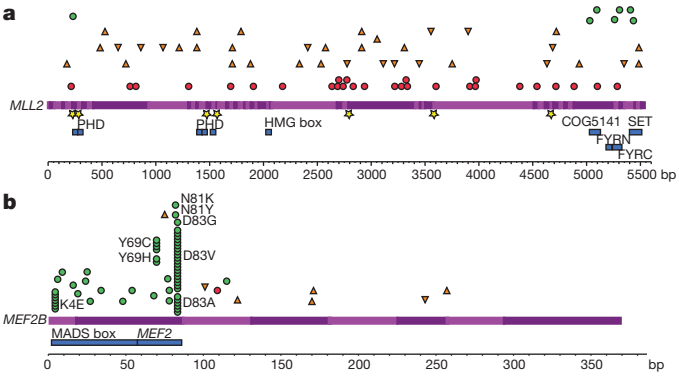



**Figure 3 | Summary and effect of somatic mutations affecting *MLL2* and *MEF2B*.** **a**, Re-sequencing the *MLL2* locus in 89 samples revealed mainly nonsense (red circles) and frameshift-inducing indel mutations (orange triangles; inverted triangles for insertions and upright triangles for deletions). A smaller number of non-synonymous somatic mutations (green circles) and point mutations or deletions affecting splice sites (yellow stars) were also observed. All of the non-synonymous point mutations affected a residue within either the catalytic SET domain, the FYRC domain (FY-rich carboxy-terminal domain) or PHD zinc finger domains. The effect of these splice-site mutations on *MLL2* splicing was also explored (Supplementary Figure 7). **b**, The cSNVs and somatic mutations found in *MEF2B* in all FL and DLBCL cases sequenced are shown with the same symbols. Only the amino acids with variants in at least two patients are labelled. cSNVs were most prevalent in the first two protein-coding exons of *MEF2B* (exons 2 and 3). The crystal structure of MEF2 bound to EP300 supports the idea that two of the mutated sites (L67 and Y69) are important in the interaction between these proteins (Supplementary Figure 8 and Supplementary Discussion)[50].

chromosome (BAC) clone sequencing in eight FL cases to show that in all eight cases the mutations were in *trans*, affecting both *MLL2* alleles. This observation is consistent with the notion that there is a complete, or near-complete, loss of *MLL2* in the tumour cells of such patients.

With the exception of two primary FL cases and two DLBCL cell lines (Pfeiffer and SU-DHL-9), the majority of *MLL2* mutations seemed to be heterozygous. Analysis of Affymetrix 500k SNP array data from two FL cases with apparent homozygous mutations revealed that both tumours showed copy number neutral loss of heterozygosity (LOH) for the region of chromosome 12 containing *MLL2* (Supplementary Methods). Thus, in addition to bi-allelic mutation, LOH is a second, albeit less common mechanism by which *MLL2* function is lost.

*MLL2* was the most frequently mutated gene in FL, and among the most frequently mutated genes in DLBCL (Fig. 2). We confirmed *MLL2* mutations in 31 of 35 FL patients (89%), in 12 of 37 DLBCL patients (32%), in 10 of 17 DLBCL cell lines (59%) and in none of the eight normal centroblast samples we sequenced. Our analysis predicted that the majority of the somatic mutations observed in *MLL2* were inactivating (91% disrupted the reading frame or were truncating point mutations), indicating to us that *MLL2* is a tumour suppressor of significance in NHL.

## Recurrent point mutations in *MEF2B*

Our selective pressure analysis also revealed genes with stronger pressure for acquisition of amino acid substitutions than for nonsense mutations. One such gene was *MEF2B*, which had not previously been linked to lymphoma. We found that 20 (15.7%) cases had *MEF2B* cSNVs and 4 (3.1%) cases had *MEF2C* cSNVs. All cSNVs detected by RNA-seq affected either the MADS box or MEF2 domains. To determine the frequency and scope of *MEF2B* mutations, we Sanger-sequenced exons 2 and 3 in 261 primary FL samples; 259 DLBCL primary tumours; 17 cell lines; 35 cases of assorted NHL (IBL, composite FL and PBMCL); and eight non-malignant centroblast samples. We also used a capture strategy (Supplementary Methods) to sequence the entire *MEF2B* coding region in the 261 FL samples, revealing six additional variants outside exons 2 and 3. We thus identified 69 cases (34 DLBCL, 12.67%; and 35 FL, 15.33%) with *MEF2B* cSNVs or indels, failing to observe novel variants in other NHL and non-malignant samples. Of the variants 55 (80%) affected residues within the MADS box and MEF2 domains encoded by exons 2 and 3 (Supplementary Table 11; Fig. 3b). Each patient generally had a single *MEF2B* variant and we observed relatively few (eight in total, 10.7%) truncation-inducing SNVs or indels. Non-synonymous SNVs were by far the most common type of change observed, with 59.4% of detected variants affecting K4, Y69, N81 or D83. In 12 cases *MEF2B* mutations were shown to be somatic, including representative mutations at each of K4, Y69, N81 and D83 (Supplementary Table 12). We did not detect mutations in ABC cases, indicating that somatic mutations in *MEF2B* have a role unique to the development of GCB DLBCL and FL (Fig. 2).

**Table 2 | Summary of types of *MLL2* somatic mutations**

| Sample Type | FL | DLBCL | DLBCL cell-line | Centroblast |
|---|---|---|---|---|
| Truncation | 18 | 4 | 7 | 0 |
| Indel with frameshift | 22 | 8 | 6 | 0 |
| Splice site | 4 | 2 | 0 | 0 |
| SNV | 3 | 2 | 2 | 0 |
| Any mutation/ number of cases | 31/35 | 12/37 | 10/17 | 0/8 |
| Percentage | 89 | 32 | 59 | 0 |



©2011 Macmillan Publishers Limited. All rights reserved

## Discussion

In our study of genome, transcriptome and exome sequences from 127 B-cell NHL cases, we identified 109 genes with clear evidence of somatic mutation in multiple individuals. Significant selection seems to act on at least 26 of these for the acquisition of either nonsense or missense mutations. To the best of our knowledge, the majority of these genes had not previously been associated with any cancer type. We observed an enrichment of somatic mutations affecting genes involved in transcriptional regulation and, more specifically, chromatin modification.

*MLL2* emerged from our analysis as a major tumour suppressor locus in NHL. It is one of six human H3K4-specific methyltransferases[29], all of which share homology with the *Drosophila* trithorax gene. Trimethylated H3K4 (H3K4me3) is an epigenetic mark associated with the promoters of actively transcribed genes. By laying down this mark, MLLs are responsible for the transcriptional regulation of developmental genes including the homeobox (*Hox*) gene family[30] which collectively control segment specificity and cell fate in the developing embryo[31,32]. Each MLL family member is thought to target different subsets of *Hox* genes[33] and in addition, MLL2 is known to regulate the transcription of a diverse set of genes[34]. Recently, *MLL2* mutations were reported in a small-cell lung cancer cell line[35] and in renal carcinoma[36], but the frequency of nonsense mutations affecting *MLL2* in these cancers was not established in these reports. Inactivating mutations were reported recently in *MLL2* or *MLL3* in 16% of medulloblastoma patients[37], further implicating *MLL2* as a cancer gene.

Our data link *MLL2* somatic mutations to B-cell NHL. The reported mutations are likely to be inactivating and in eight of the cases with multiple mutations, we confirmed that both alleles were affected, presumably resulting in essentially complete loss of MLL2 function. The high prevalence of *MLL2* mutations in FL (89%) equals the frequency of the t(14;18)(q32;q21) translocation, which is considered the most prevalent genetic abnormality in FL[3]. In DLBCL tumour samples and cell lines, *MLL2* mutation frequencies were 32% and 59%, respectively, also exceeding the prevalence of the most frequent cytogenetic abnormalities, such as the various translocations involving 3q27, which occur in 25–30% of DLBCLs and are enriched in ABC cases[38]. Importantly, we found *MLL2* mutated in both DLBCL subtypes (Fig. 2). Our analyses thus indicate that *MLL2* acts as a central tumour suppressor in FL and both DLBCL subtypes.

The *MEF2* gene family encodes four related transcription factors that recruit histone-modifying enzymes including histone deacetylases (HDACs) and HATs in a calcium-regulated manner. Although truncating variants were detected in our analysis of *MEF2* gene family members, our analysis suggests that, in contrast to *MLL2*, *MEF2* family members tend to selectively acquire non-synonymous amino acid substitutions. In the case of *MEF2B*, 59.4% of all the cSNVs were found at four sites within the protein (K4, Y69, N81 and D83), and all four of these sites were confirmed to be targets of somatic mutation. D83 is affected in 39% of the *MEF2B* alterations, resulting in replacement of the charged aspartate with any of alanine, glycine or valine. Although we cannot yet predict the consequences of these substitutions on protein function, it seems likely that their effect would have an impact on the ability of MEF2B to facilitate gene expression and thus have a role in promoting the malignant transformation of germinal centre B cells to lymphoma (Supplementary Discussion).

*MEF2B* mutations can be linked to *CREBBP* and *EP300* mutations, and to recurrent Y641 mutations in *EZH2* (ref. 13). One target of CREBBP/EP300 HAT activity is H3K27, which is methylated by EZH2 to repress transcription. There is evidence that the action of EZH2 antagonizes that of CREBBP/EP300 (ref. 39). One function of MEF2 is to recruit either HDACs or CREBBP/EP300 to target genes[40], and it has been suggested that HDACs compete with CREBBP/EP300 for the same binding site on MEF2 (ref. 41). Under normal $Ca^{2+}$ levels, MEF2 is bound by type IIa HDACs, which maintain the tails of histone proteins in a deacetylated repressive chromatin state[42]. Increased cytoplasmic $Ca^{2+}$ levels induce the nuclear export of HDACs, enabling the recruitment of HATs such as CREBBP/EP300, facilitating transcription at MEF2 target genes. Mutation of *CREBBP*, *EP300* or *MEF2B* may have an impact on the expression of MEF2 target genes owing to reduced acetylation of nucleosomes near these genes (Supplementary Figure 5; Supplementary Discussion). In light of the recent finding that heterozygous *EZH2* Y641 mutations enhance overall H3K27 trimethylation activity of PRC2 (refs 43, 44), it is possible that mutation of both *MLL2* and *EZH2* could cooperate in reducing the expression of some of the same target genes. Our data indicate that (1) post-transcriptional modification of histones is of key importance in germinal centre B cells and (2) deregulated histone modification due to these mutations is likely to result in reduced acetylation and enhanced methylation, and acts as a core driver event in the development of NHL (Supplementary Figure 5).

## METHODS SUMMARY

All samples analysed contained at least 50% tumour cells. Genomes, exomes and transcriptomes were sequenced using a combination of Illumina GAIIx and HiSeq 2000 instruments to read lengths of between 36 and 100 nucleotides. Exome capture was performed using the Agilent SureSelect Target Enrichment System Protocol (Version 1.0, September 2009). Alignment was accomplished using BWA[45] and variants were identified using SNVmix[46]. Variants were manually reviewed in IGV and were confirmed (where applicable) by PCR followed by either Sanger sequencing or Illumina re-sequencing. Structural rearrangements in genomes and transcriptomes were identified using ABySS[47]. Gene expression values used for subtype assignment were calculated as reads per kilobase gene model per million mapped reads (RPKM) values[48] and subtypes were assigned using an adaptation of the method developed for data from Affymetrix expression arrays[49] trained with samples previously classified by this standard approach.

**Received 13 November 2010; accepted 7 July 2011.**
**Published online 27 July 2011.**

1. Anderson, J. R., Armitage, J. O., Weisenburger, D. D., Non-Hodgkin's Lymphoma Classification Project. Epidemiology of the non-Hodgkin's lymphomas: distributions of the major subtypes differ by geographic locations. *Ann. Oncol.* **9,** 717–720 (1998).
2. Lenz, G. & Staudt, L. M. Aggressive lymphomas. *N. Engl. J. Med.* **362,** 1417–1429 (2010).
3. Horsman, D. E. *et al.* Follicular lymphoma lacking the t(14;18)(q32;q21): identification of two disease subtypes. *Br. J. Haematol.* **120,** 424–433 (2003).
4. Iqbal, J. *et al. BCL2* translocation defines a unique tumor subset within the germinal center B-cell-like diffuse large B-cell lymphoma. *Am. J. Pathol.* **165,** 159–166 (2004).
5. Lenz, G. *et al.* Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc. Natl Acad. Sci. USA* **105,** 13520–13525 (2008).
6. Pasqualucci, L. *et al.* Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. *J. Exp. Med.* **203,** 311–317 (2006).
7. Kato, M. *et al.* Frequent inactivation of A20 in B-cell lymphomas. *Nature* **459,** 712–716 (2009).
8. Compagno, M. *et al.* Mutations of multiple genes cause deregulation of NF-κB in diffuse large B-cell lymphoma. *Nature* **459,** 717–721 (2009).
9. Davis, R. E. *et al.* Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* **463,** 88–92 (2010).
10. Ngo, V. N. *et al.* Oncogenically active *MYD88* mutations in human lymphoma. *Nature* **470,** 115–119 (2011).
11. Mardis, E. R. *et al.* Recurring mutations found by sequencing an acute myeloid leukemia genome. *N. Engl. J. Med.* **361,** 1058–1066 (2009).
12. Shah, S. P. *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461,** 809–813 (2009).
13. Morin, R. D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nature Genet.* **42,** 181–185 (2010).
14. Futreal, P. A. *et al.* A census of human cancer genes. *Nature Rev. Cancer* **4,** 177–183 (2004).
15. Pasqualucci, L. *et al.* Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* **471,** 189–195 (2011).
16. Yusuf, I., Zhu, X., Kharas, M. G., Chen, J. & Fruman, D. A. Optimal B-cell proliferation requires phosphoinositide 3-kinase-dependent inactivation of FOXO transcription factors. *Blood* **104,** 784–787 (2004).
17. Saito, M. *et al.* BCL6 suppression of BCL2 via Miz1 and its disruption in diffuse large B cell lymphoma. *Proc. Natl Acad. Sci. USA* **106,** 11294–11299 (2009).
18. Lenz, G. *et al.* Oncogenic *CARD11* mutations in human diffuse large B cell lymphoma. *Science* **319,** 1676–1679 (2008).



©2011 Macmillan Publishers Limited. All rights reserved

19. Greenman, C., Wooster, R., Futreal, P. A., Stratton, M. R. & Easton, D. F. Statistical analysis of pathogenicity of somatic mutations in cancer. *Genetics* **173,** 2187–2198 (2006).
20. Cheung, K. J. *et al.* Acquired *TNFRSF14* mutations in follicular lymphoma are associated with worse prognosis. *Cancer Res.* **70,** 9166–9174 (2010).
21. Du, M. Q. *et al. BCL10* gene mutation in lymphoma. *Blood* **95,** 3885–3890 (2000).
22. Kreutz, B., Hajicek, N., Yau, D. M., Nakamura, S. & Kozasa, T. Distinct regions of Gα13 participate in its regulatory interactions with RGS homology domain-containing RhoGEFs. *Cell. Signal.* **19,** 1681–1689 (2007).
23. Bhattacharyya, R. & Wedegaertner, P. Gα13 requires palmitoylation for plasma membrane localization, Rho-dependent signaling, and promotion of p115-RhoGEF membrane binding. *J. Biol. Chem.* **275,** 14992–14999 (2000).
24. Manganello, J. M., Huang, J., Kozasa, T., Voyno-Yasenetskaya, T. A. & Le Breton, G. C. Protein kinase A-mediated phosphorylation of the Gα13 switch I region alters the Gαβγ13-G protein-coupled receptor complex and inhibits Rho activation. *J. Biol. Chem.* **278,** 124–130 (2003).
25. Brunet, A. *et al.* Protein kinase SGK mediates survival signals by phosphorylating the forkhead transcription factor FKHRL1 (FOXO3a). *Mol. Cell. Biol.* **21,** 952–965 (2001).
26. Tai, D. J. C., Su, C.-C., Ma, Y.-L. & Lee, E. H. Y. SGK1 phosphorylation of IκB kinase α and p300 Up-regulates NF-κB activity and increases *N*-methyl-D-aspartate receptor NR2A and NR2B expression. *J. Biol. Chem.* **284,** 4073–4089 (2009).
27. Mo, J. *et al.* Serum- and glucocorticoid-inducible kinase 1 (SGK1) controls Notch1 signaling by downregulation of protein stability through Fbw7 ubiquitin ligase. *J. Cell Sci.* **124,** 100–112 (2011).
28. Young, K. H. *et al.* Structural profiles of *TP53* gene mutations predict clinical outcome in diffuse large B-cell lymphoma: an international collaborative study. *Blood* **112,** 3088–3098 (2008).
29. Shilatifard, A. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr. Opin. Cell Biol.* **20,** 341–348 (2008).
30. Milne, T. *et al.* MLL targets SET domain methyltransferase activity to *Hox* gene promoters. *Mol. Cell* **10,** 1107–1117 (2002).
31. Krumlauf, R. *Hox* genes in vertebrate development. *Cell* **78,** 191–201 (1994).
32. Canaani, E. *et al.* ALL-1//MLL1, a homologue of *Drosophila* TRITHORAX, modifies chromatin and is directly involved in infant acute leukaemia. *Br. J. Cancer* **90,** 756–760 (2004).
33. Wang, P. *et al.* Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Mol. Cell. Biol.* **29,** 6074–6085 (2009).
34. Issaeva, I. *et al.* Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. *Mol. Cell. Biol.* **27,** 1889–1903 (2007).
35. Pleasance, E. D. *et al.* A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463,** 184–190 (2010).
36. Dalgliesh, G. L. *et al.* Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463,** 360–363 (2010).
37. Parsons, D. W. *et al.* The genetic landscape of the childhood cancer medulloblastoma. *Science* **331,** 435–439 (2011).
38. Iqbal, J. *et al.* Distinctive patterns of BCL6 molecular alterations and their functional consequences in different subgroups of diffuse large B-cell lymphoma. *Leukemia* **21,** 2332–2343 (2007).
39. Pasini, D. *et al.* Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res.* (2010).
40. Giordano, A. & Avantaggiati, M. p300 and CBP: partners for life and death. *J. Cell. Physiol.* **181,** 218–230 (1999).
41. Han, A., He, J., Wu, Y., Liu, J. O. & Chen, L. Mechanism of recruitment of class II histone deacetylases by myocyte enhancer factor-2. *J. Mol. Biol.* **345,** 91–102 (2005).
42. Youn, H. & Liu, J. Cabin1 represses MEF2-dependent Nur77 expression and T cell apoptosis by controlling association of histone deacetylases and acetylases with MEF2. *Immunity* **13,** 85–94 (2000).
43. Yap, D. B. *et al.* Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. *Blood* **117,** 2451–2459 (2011).
44. Sneeringer, C. J. *et al.* Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proc. Natl Acad. Sci. USA* **107,** 20980–20985 (2010).
45. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).
46. Goya, R. *et al.* SNVMix: predicting single nucleotide variants from next-generation sequencing of tumors. *Bioinformatics* **26,** 730–736 (2010).
47. Robertson, G. *et al. De novo* assembly and analysis of RNA-seq data. *Nature Methods* **7,** 909–912 (2010).
48. Mortazavi, A., Williams, B. A., Mccue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5,** 621–628 (2008).
49. Wright, G. *et al.* A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc. Natl Acad. Sci. USA* **100,** 9991–9996 (2003).
50. He, J. *et al.* Structure of p300 bound to MEF2 on DNA reveals a mechanism of enhanceosome assembly. *Nucleic Acids Res.* (2011).


**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** This study was funded in part by funding from the National Cancer Institute Office of Cancer Genomics (Contract No. HHSN261200800001E), the Terry Fox Foundation (grant 019001, Biology of Cancer: Insights from Genomic Analyses of Lymphoid Neoplasms) and Genome Canada/Genome British Columbia Grant Competition III (Project Title: High Resolution Analysis of Follicular Lymphoma Genomes) to J.M.C., R.D.G. and M.A.M. We acknowledge support from NIH grants P50CA130805-01 ''SPORE in Lymphoma, Tissue Resource Core (PI Fisher)'' and 1U01CA114778 ''Molecular Signatures to Improve Diagnosis and Outcome in Lymphoma (PI Chan)''. A.J.M. is a Career Development Program Fellow of the Leukemia and Lymphoma Society. N.A.J. was a research fellow of the Terry Fox Foundation (award NCIC 019005) and the Michael Smith Foundation for Health Research (ST-PDF-01793). M.A.M. is a Terry Fox Young Investigator and a Michael Smith Senior Research Scholar. R.D.M. is a Vanier Scholar (CIHR) and holds a MSFHR senior graduate studentship. M.M.-L. acknowledges support from a Postdoctoral Fellowship from the Spanish Ministry of Education, under the ''Programa Nacional de Movilidad de Recursos Humanos del Plan Nacional de I-D+i 2008-2011''. D.W.S. was supported by the Terry Fox Foundation Strategic Health Research Training Program in Cancer Research at Canadian Institutes of Health Research (Grant No. TGT-53912). J.J.S. acknowledges funding from The Canadian Cancer Society and the Canadian Institutes of Health Research. R.G. is supported by a UBC Four Year Fellowship. I.M.M. acknowledges the Canadian Foundation for Innovation for a Leaders Opportunity Fund. The laboratory work for this study was undertaken at the Genome Sciences Centre, British Columbia Cancer Research Centre and the Centre for Translational and Applied Genomics, a program of the Provincial Health Services Authority Laboratories. The authors would like to thank C. Greenman for supplying his software and also acknowledge D. Gerhard and S. Aparicio for discussions and guidance. Special thanks to C. Suragh, R. Roscoe, A. Troussard and A. Drobnies for expert project management assistance, and to the Library Construction, Sequencing and Bioinformatics teams at the Genome Sciences Centre. The content of this publication does not necessarily reflect the views of policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

**Author Contributions** M.A.M., R.D.G., D.E.H., M.H. and J.M.C. conceived of the study and led the design of the experiments. R.D.M. performed the analysis of sequence data, identified mutations and, with M.M.-L., A.J.M. and M.A.M., produced figures and wrote the manuscript. M.M.-L., A.J.M., D.L.T., S.Chan, S.Chittarajan, D.S., H.M., J.S., M.M., T.Z., A.D., K.T., Y.B., M.R.F., J.T.-W. and T.M.S. designed and performed experiments to amplify, discover and validate mutations. R.G., M.G. and I.M.M. contributed to analyses and reviewed the manuscript. N.A.J., M.B., B.W. and B.M. prepared the samples, performed sample sorting and COO analysis and contributed to the text. A.B.-W. and J.J.S. collected and prepared constitutional DNA samples. K.L.M., R.C., S.L., M.F. and S.J. generated *de novo* assemblies and identified mutations. M.K., S.R., M.G., O.Y. and E.Y.Z. wrote software and contributed to figures. R.D.C. performed copy number analysis and produced a figure and S.B.-N. performed confirmatory FISH experiments. Y.Z. and A.T. produced the sequencing libraries. I.B., R.H., S.J.M.J., R.M., J.S. and M.H. contributed to the development of experimental and analytical protocols. L.R. provided materials and reviewed the manuscript.

**Author Information** The SRA accession number for the submission of the data not included in previous publications is SRP001599, which is linked to the dbGAP study accession phs000235.v2.p1. Reprints and permissions information is available at www.nature.com/reprints. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at www.nature.com/nature. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.A.M. (mmarra@bcgsc.ca).



©2011 **Macmillan Publishers Limited. All rights reserved**

LETTER

doi:10.1038/nature10320

# Central powering of the largest Lyman-α nebula is revealed by polarized radiation

Matthew Hayes[1,2,3], Claudia Scarlata[4,5] & Brian Siana[6]

**High-redshift Lyman-α (Lyα) blobs[1,2] are extended, luminous but rare structures that seem to be associated with the highest peaks in the matter density of the Universe[3–6]. Their energy output and morphology are similar to those of powerful radio galaxies[7], but the source of the luminosity is unclear. Some blobs are associated with ultraviolet or infrared bright galaxies, suggesting an extreme starburst event or accretion onto a central black hole[8–10]. Another possibility is gas that is shock-excited by supernovae[11,12]. But not all blobs are associated with galaxies[13,14], and these ones may instead be heated by gas falling into a dark-matter halo[15–19]. The polarization of the Lyα emission can in principle distinguish between these options[20–22], but a previous attempt to detect this signature returned a null detection[23]. Here we report observations of polarized Lyα from the blob LAB1 (ref. 2). Although the central region shows no measurable polarization, the polarized fraction (*P*) increases to ~20 per cent at a radius of 45 kiloparsecs, forming an almost complete polarized ring. The detection of polarized radiation is inconsistent with the *in situ* production of Lyα photons, and we conclude that they must have been produced in the galaxies hosted within the nebula, and re-scattered by neutral hydrogen.**

The Lyα emission line of neutral hydrogen is a frequently used observational tracer of evolving galaxies in the high-redshift Universe. Lyα imaging surveys typically find a large number of faint unresolved objects and a small fraction of extremely luminous and spatially extended systems that are usually referred to independently as Lyman-α blobs (LABs). The compact sources usually seem to be more ordinary star-forming galaxies whereas, since their discovery, much controversy has surrounded the true nature of LABs. Because one of the possible modes of powering LABs is the accretion of gas from the intergalactic medium[15–19], an eventual consensus about this powering mechanism holds particular relevance for how the most massive and rapidly evolving high-redshift dark-matter haloes acquire the gas required to fuel star formation in the galaxies they host. Predictions for polarized Lyα radiation were made more than a decade ago[20], and have been further developed to the extent that the polarization signal may act as a diagnostic to study the origin of the Lyα photons[21]. By relying on the scattering of Lyα photons at large distances from their production sites, these polarization measurements may also provide an independent constraint on the distribution of neutral gas in the circumgalactic regions, which has proven a difficult measurement to make by traditional techniques. The one previous attempt to detect the polarization of a LAB returned a null result[23], although very deep observations are required that need considerable investments of time on 8–10-m-class telescopes. In the observations made with the European Southern Observatory's Very Large Telescope presented here, we observe one of the largest and most luminous LABs on the sky in imaging polarimetry mode (Figs 1 and 2).

These new, deep polarimetric observations show that a substantial fraction of the Lyα emission from LAB1 is polarized. The data enable the fraction of light that is linearly polarized (*P*) to be measured to better than 5% (root mean squared) in the regions of peak surface brightness; however, in these central regions we find essentially no polarization signal. To detect the polarized emission we need to examine the fainter signal emanating a few arcseconds from the centre—when Lyα is emitted with a high polarization fraction ($P \approx 20\%$), it is not spatially coincident with the brighter regions of the nebula. Indeed, the highest and most statistically certain measures of *P* are found some 4–8 arcsec from the centre, forming an almost complete ring of polarized emission. Just a single resolution element shows a signal-to-noise ratio on *P* that exceeds 3, but our confidence in these results is substantially enhanced by the strong degree of spatial correlation between the resolution elements that are observed to be polarized at high significance. Away from the regions of bright Lyα emission, our data show no measurable polarization signal, but a second nearby extended Lyα source (LAB8; ref. 4) that lies to the north also shows strongly polarized Lyα emission.

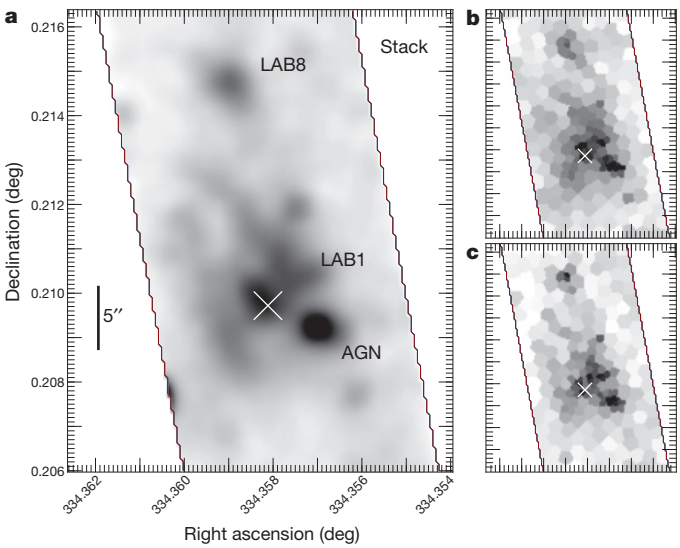



**Figure 1 | LAB1 in Lyα. a**, Combined intensity frame that has been adaptively smoothed to show detail. The Lyα emission is spatially extended over scales of more than 10 arcsec (80 kpc). The 5-arcsecond scale bar corresponds to 40 physical kiloparsecs. North is up and east is to the left. The primary target of this observation is LAB1 (ref. 2), although the position angle was chosen so as to also include LAB8 (ref. 4), which resides about 20 arcsec to the North. The bright point source to the southeast (lower right) is the active galactic nucleus (AGN) C11 (ref. 2). Major centres are labelled in the image. **b**, **c**, The ordinary (**c**) and extraordinary (**b**) beams, tessellated for quantitative analysis (Supplementary Information). All images are shown on the same intensity scaling, and even by eye there are visible differences in the intensity structure between the two beams. The cross marks the point that we determine to be the centre in the subsequent analysis, and is determined as the point of peak Lyα surface brightness that is not associated with the AGN.

[1]Université de Toulouse, UPS-OMP, IRAP, Toulouse, France. [2]CNRS, IRAP, 14 avenue Edouard Belin, F-31400 Toulouse, France. [3]Observatory of Geneva, University of Geneva, 51 chemin des Maillettes, 1290 Versoix, Switzerland. [4]Minnesota Institute for Astrophysics, School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota 55455, USA. [5]Spitzer Science Center, California Institute of Technology, 220-6, Pasadena, California 91125, USA. [6]Department of Astronomy, California Institute of Technology, MS 249-17, Pasadena, California 91125, USA.



©2011 Macmillan Publishers Limited. All rights reserved

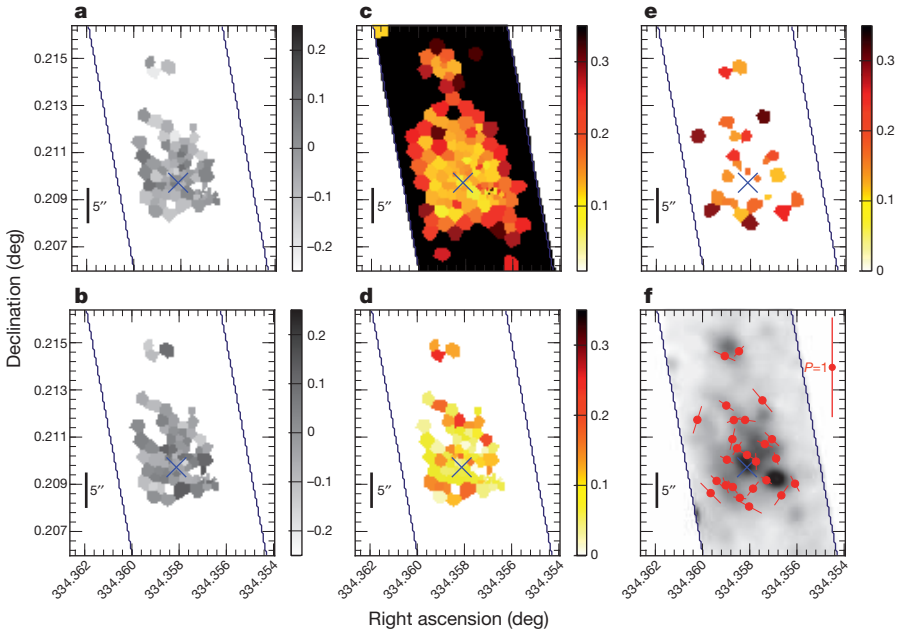



**Figure 2 | The polarization of LAB1.** Scaling is shown to the right of each panel. **a**, **b**, The $Q$ (**a**) and $U$ (**b**) Stokes parameters. **c**, Map of the polarization fraction $P$ that we can measure at 95.4% confidence; this error map demonstrates that in individual resolution elements we can constrain $P$ in the central regions down to 10% ($2\sigma$). At a radius of ~6–7 arcsec, this decreases to about the 20% level, and beyond ~10 arcsec $P$ becomes noise-dominated. **d**, Value of $P$ in resolution elements where it can be well constrained; we make a cut based on the error map, and show bins where $P$ can be measured to better than 16% ($2\sigma$). $P \approx$ 5–10% is found in the centre, but increases to 15–20% at ~5 arcsec from the centre towards the south, west and north edges of LAB1. $P \approx$ 20% is also seen around LAB8. **e**, $P$ in bins for which we can be at least 95.4% confident of its measurement. The structure around LAB1 and LAB8 remains, extending to larger radii and forming a broken ring at $P \approx$ 20%. **f**, A smoothed intensity image with the direction of the polarization vector $\chi$ shown by red bars. As in **e**, only bins showing measurements at more than 95.4% confidence level are shown. There is a clear trend that, when bins show strong polarization, $\chi$ orientates itself tangentially to the overall structure and local Ly$\alpha$ surface brightness. These features conform well with the predictions for the polarization of Ly$\alpha$ photons that are centrally produced and scatter at large radii, or are produced by cooling gas that flows onto the dark-matter halo in quasi-spherical symmetry[21].

Ly$\alpha$ is a resonant line, and photons scatter in neutral hydrogen (H I), undergoing a complicated radiation transfer. At each scattering event, of which there may be millions, a photon acquires a new direction and frequency, depending on the kinematics of the scattering medium and quantum-mechanical probabilities[24]. The outgoing frequency determines the optical depth that the photon then sees[25,26]: small shifts from line-centre (core scatterings) result in effectively no transfer, whereas long-frequency excursions (wing scatterings) result in significant flights and likely escape from the system. The more frequent core scatterings may emit polarized radiation, but it is typically with a low fraction ($P \lesssim 7\%$); in contrast, the comparatively rare wing scatterings may introduce a high degree of linear polarization to the outgoing radiation ($P$ up to 40%)[21,24]. The observed polarization signal therefore provides an unambiguous signature of the scattering nature of the diffuse Ly$\alpha$ halos and the distribution of circumgalactic gas. At the atomic level, the detected polarization signal provides evidence that the Ly$\alpha$ photons are not emerging after a large number of resonant core scattering events but are leaving after long shifts in frequency.

The detection of polarization demonstrates that the Ly$\alpha$ photons cannot have been produced *in situ* at large radial distances. To impart a significant level of polarization on the outgoing radiation, a geometry is required in which Ly$\alpha$ photons preferentially escape after scattering off hydrogen atoms at angles of 90° to their vector before the scattering event[21,24]. This geometry is not expected in the case of wind-induced shock excitation because thermalized cooling gas will generate Ly$\alpha$ photons with no preferential orientation or impact angle with respect to the scattering medium. Similarly, the filamentary streams of cold infalling gas predicted by simulation[27,28] seem to have volume filling factors that are too small, and to be too narrow (a few kiloparsecs) for any imparted polarization signal to be observable[17]—a spatial resolution would be required that is currently unattainable by optical polarimeters. Although neither of these rejected scenarios has yet been thoroughly tested by radiative transport studies, the production of polarized Ly$\alpha$ is well understood, and it is difficult to see a mechanism by which observably polarized Ly$\alpha$ may be produced in these situations. We therefore conclude that the Ly$\alpha$ photons must have been produced centrally in the galaxies themselves and scattered at large radii by neutral gas. Even if the halo is accreting intergalactic gas, this process is not responsible for a large fraction of the Ly$\alpha$ emission.

The geometrical origin of the polarization signal in the photon scattering model, which is analogous to Rayleigh scattering, implies also that the angle of the polarization vector $\chi$ should align tangentially to the overall geometry of the system at large radii[21,22]. We also extract this $\chi$ vector from the data (Fig. 2f). If we define the geometric centre to be the region of peak Ly$\alpha$ surface brightness (excluding the active nucleus; Fig. 1), these vectors show a strong tendency to orient along directions tangential to the overall geometry of the system. Again, this is in good agreement with theoretical expectations, should the Ly$\alpha$ photons be produced centrally by ordinary star-formation and/or active-galactic-nucleus-fuelled processes and scatter at large radii[21]. However, unlike the idealized models used for radiative transport simulations, which assume perfect spherical symmetry and no velocity shear or clumping of the gas, the true Ly$\alpha$ surface brightness is far from smooth and symmetric. This may reflect either multiple Ly$\alpha$ sources within the halo or large-scale inhomogeneities in the scattering medium, or both. Polarization of Ly$\alpha$ emission in a clumpy medium has not yet been tested by simulation, although clumping could potentially either suppress or enhance the polarization[21], and probably depends on the relative size and volume-filling factor of the clumps. In any event, we would still expect the polarization signature to remain, provided that the individual regions are well resolved, in which case $\chi$ would then be expected to align tangentially to contours in the local surface brightness. This effect is also reflected in the data, and we find that polarization vectors line up preferentially at angles roughly



©2011 Macmillan Publishers Limited. All rights reserved

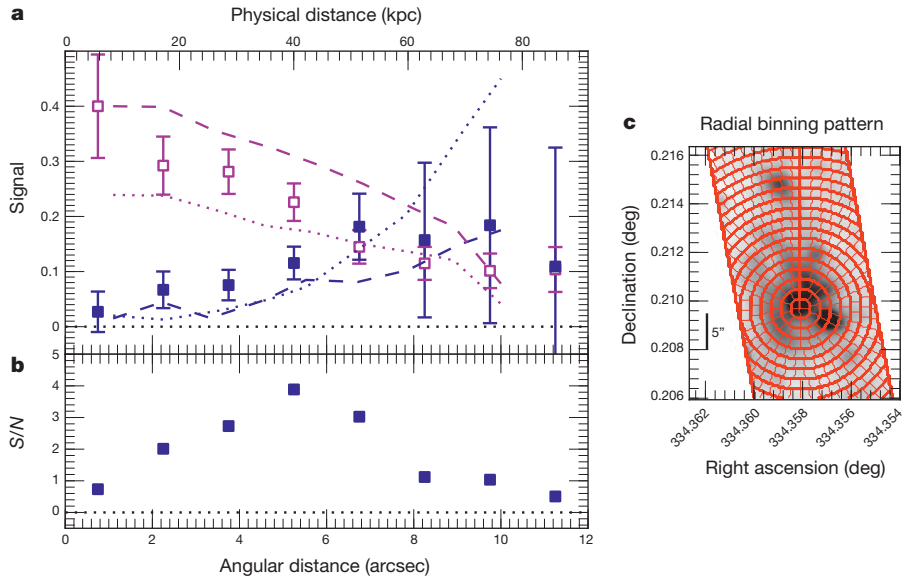



**Figure 3 | The radial polarization profile.** **a**, Blue points show the fractional polarization, *P*, which is measured in concentric 1.5 arcsec annuli and corrected for spuriously high polarization fractions that may be measured at low signal-to-noise ratio[29]. The total intensity profile, *I*, is shown in magenta. Error bars on both *P* and *I* are $1\sigma$. **b**, The signal-to-noise (*S*/*N*) ratio of *P*. In the central regions where intensity is highest, *P* is measured at ~4%, although at $1\sigma$ it is consistent with no polarization. *P* and its signal-to-noise ratio both increase with radius; by 5 arcsec $P = 10\%$ and is inconsistent with zero at the $4\sigma$ level. *P* peaks at 18% at 7 arcsec, beyond which radius the noise becomes dominant. Overlaid on **a** are the results of numerical models for scattered Ly$\alpha$ radiation[21] in an expanding H I shell with two column densities: $10^{19}$ cm$^{-2}$ (dashed) and $10^{20}$ cm$^{-2}$ (dotted). These curves have been smoothed over lengths of 1.5 arcsec to match our annuli. The actual radius of the shell is unknown, so the theoretical points have been scaled to a shell radius of 10 arcsec, to match both the surface brightness and polarization profiles approximately. This shows that the predictions are qualitatively matched by observations. Models for scattering in a static intergalactic medium and spherically symmetric accretion are not shown because they are not leading explanations for the powering of LABs. **c**, The different binning method (Supplementary Information) that enables annular measurements of *P* using independent bins. The centre is the same as illustrated in Fig. 1.

perpendicular to the local surface brightness gradient at their position. The Supplementary Information includes a quantitative analysis of the polarization angles compared with both local gradients in surface brightness and the overall system geometry.

There is one final observable prediction that has been suggested by radiative transport simulations. This is the radial profile of the polarization fraction *P*, which according to the scattering shell model should manifest as with $P \approx 0$ in the geometric centre and increase with increasing radius. We have computed this radial profile of *P* using the same central coordinates as defined previously (Fig. 3) and find strong evidence for *P* increasing with radius, again in agreement with predictions. The central point on the profile and the most significantly polarized point in the ring are discrepant at over $2\sigma$. We have fitted the innermost 7 arcsec (five data points; where data are deemed reliable) with the simplest function for *P* varying linearly with radius. Computing the classical *p* value, $1 - \Gamma(\nu/2,\chi^2/2)$, where $\nu$ is the number of degrees of freedom, shows more than 96% confidence in this hypothesis. The same test performed on a flat profile ($P = 8.3\%$) suggests that a structure in which *P* is independent of radius is consistent with the data at 16%. In any event, the slim possibility of *P* that is flat with radius does not present a difficulty for the scattering model, which is supported not only by the detection of polarized Ly$\alpha$ itself but also by the angular distribution of polarization vectors; this radial profile is simply the least robust of our constraints. The actual gradient of *P* with radius depends on both the column density and radius of the shell, neither of which we know—assuming that the idealized geometry holds and that the shell has a radius of 10 arcsec, the models with column densities of $10^{19}$–$10^{20}$ cm$^{-2}$ would be in agreement with the data (Fig. 3). In this step of the analysis we also compute the total luminosity-weighted polarization fraction within each radius: within 7 arcsec we obtain an integrated value of $P = 11.9 \pm 2\%$. We take this value as the overall polarization fraction of the nebula, which we quote at $6\sigma$ significance.

**Received 2 April; accepted 20 June 2011.**


1. Francis, P. J. *et al.* A group of galaxies at redshift 2.38. *Astrophys. J.* **457,** 490–499 (1996).
2. Steidel, C. C. *et al.* Ly$\alpha$ imaging of a proto-cluster region at $<z> = 3.09$. *Astrophys. J.* **532,** 170–182 (2000).
3. Palunas, P., Teplitz, H. I., Francis, P. J., Williger, G. M. & Woodgate, B. E. The distribution of Ly$\alpha$-emitting galaxies at $z = 2.38$. *Astrophys. J.* **602,** 545–554 (2004).
4. Matsuda, Y. *et al.* A Subaru search for Ly$\alpha$ blobs in and around the protocluster region at redshift $z = 3.1$. *Astron. J.* **128,** 569–584 (2004).
5. Yang, Y., Zabludoff, A., Tremonti, C., Eisenstein, D. & Davé, R. Extended Ly$\alpha$ nebulae at $z \approx 2.3$: an extremely rare and strongly clustered population? *Astrophys. J.* **693,** 1579–1587 (2009).
6. Prescott, M. K. M., Kashikawa, N., Dey, A. & Matsuda, Y. The overdense environment of a large Ly$\alpha$ nebula at $z \sim 2.7$. *Astrophys. J.* **678,** L77–L80 (2008).
7. Saito, T. *et al.* Systematic survey of extended Ly$\alpha$ sources over $z \sim 3$–5. *Astrophys. J.* **648,** 54–66 (2006).
8. Geach, J. E. *et al.* The Chandra Deep Protocluster Survey: Ly$\alpha$ blobs are powered by heating, not cooling. *Astrophys. J.* **700,** 1–9 (2009).
9. Scarlata, C. *et al.* He II emission in Ly$\alpha$ nebulae: active galactic nucleus or cooling radiation? *Astrophys. J.* **706,** 1241–1252 (2009).
10. Colbert, J. W. *et al.* Polycyclic aromatic hydrocarbon emission within Ly$\alpha$ blobs. *Astrophys. J.* **728,** 59–71 (2011).
11. Taniguchi, Y. & Shioya, Y. Superwind model of extended Ly$\alpha$ emitters at high redshift. *Astrophys. J.* **532,** L13–L16 (2000).
12. Mori, M., Umemura, M. & Ferrara, A. The nature of Ly$\alpha$ blobs: supernova-dominated primordial galaxies. *Astrophys. J.* **613,** L97–L100 (2004).
13. Nilsson, K. K., Fynbo, J. P. U., Møller, P., Sommer-Larsen, J. & Ledoux, C. A. Lyman-$\alpha$ blob in the GOODS South field: evidence for cold accretion onto a dark matter halo. *Astron. Astrophys.* **452,** L23–L26 (2006).
14. Smith, D. J. B. & Jarvis, M. J. Evidence for cold accretion onto a massive galaxy at high redshift? *Mon. Not. R. Astron. Soc.* **378,** L49–L53 (2007).
15. Fardal, M. A. *et al.* Cooling radiation and the Ly$\alpha$ luminosity of forming galaxies. *Astrophys. J.* **562,** 605–617 (2001).
16. Yang, Y. *et al.* Probing galaxy formation with He II cooling lines. *Astrophys. J.* **640,** 539–552 (2006).
17. Dijkstra, M. & Loeb, A. Ly$\alpha$ blobs as an observational signature of cold accretion streams into galaxies. *Mon. Not. R. Astron. Soc.* **400,** 1109–1120 (2009).
18. Goerdt, T. *et al.* Gravity-driven Ly$\alpha$ blobs from cold streams into galaxies. *Mon. Not. R. Astron. Soc.* **407,** 613–631 (2010).
19. Faucher-Giguère, C.-A., Kereš, D., Dijkstra, M., Hernquist, L. & Zaldarriaga, M. Ly$\alpha$ cooling emission from galaxy formation. *Astrophys. J.* **725,** 633–657 (2010).




©2011 Macmillan Publishers Limited. All rights reserved

20. Lee, H. & Ahn, S. Polarization of the Ly α from an anisotrophy expanding H I shell in primeval galaxies. *Astrophys. J.* **504,** L61–L64 (1998).
21. Dijkstra, M. & Loeb, A. The polarization of scattered Lyα radiation around high-redshift galaxies. *Mon. Not. R. Astron. Soc.* **386,** 492–504 (2008).
22. Rybicki, G. B. & Loeb, A. Polarization of the Lyα halos around sources before cosmological reionization. *Astrophys. J.* **520,** L79–L81 (1999).
23. Prescott, M. K. M., Smith, P. S., Schmidt, G. D. & Dey, A. The line polarization within a giant Lyα nebula. *Astrophys. J.* **730,** L25–L30 (2011).
24. Stenflo, J. O. Resonance-line polarization. V. Quantum-mechanical interference between states of different total angular momentum. *Astron. Astrophys.* **84,** 68–74 (1980).
25. Adams, T. F. The escape of resonance-line radiation from extremely opaque media. *Astrophys. J.* **174,** 439–448 (1972).
26. Neufeld, D. A. The transfer of resonance-line radiation in static astrophysical media. *Astrophys. J.* **350,** 216–241 (1990).
27. Kereš, D., Katz, N., Weinberg, D. H. & Davé, R. How do galaxies get their gas? *Mon. Not. R. Astron. Soc.* **363,** 2–28 (2005).
28. Dekel, A. & Birnboim, Y. Gravitational quenching in massive galaxies and clusters by clumpy accretion. *Mon. Not. R. Astron. Soc.* **383,** 119–138 (2008).
29. Wardle, J. F. C. & Kronberg, P. P. The linear polarization of quasi-stellar radio sources at 3.71 and 11.1 centimeters. *Astrophys. J.* **194,** 249–255 (1974).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank P. Ogle and J. Colbert for valuable suggestions about polarimetric observations; N. Panagia and T. Jones for comments on the manuscript; and D. Schaerer, N. Scoville, C. Lidman, A. Dey, M. Prescott and P. Lynam for discussions. This work was based on observations made with European Southern Observatory telescopes at the Paranal Observatory under programme ID 084.A-0954. M.H. was supported in part by the Swiss National Science Foundation, and also received support from Agence Nationale de la Recherche (reference ANR-09-BLAN-0234-01).

**Author Contributions** All authors contributed to the proposal preparation. M.H. and C.S. observed and reduced the data. M.H. analysed the results. All authors contributed to the manuscript preparation.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.H. (matthew.hayes@ast.obs-mip.fr).



©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10344

# Suppression of the coffee-ring effect by shape-dependent capillary interactions

Peter J. Yunker[1], Tim Still[1,2], Matthew A. Lohr[1] & A. G. Yodh[1]

**When a drop of liquid dries on a solid surface, its suspended particulate matter is deposited in ring-like fashion. This phenomenon, known as the coffee-ring effect[1–3], is familiar to anyone who has observed a drop of coffee dry. During the drying process, drop edges become pinned to the substrate, and capillary flow outward from the centre of the drop brings suspended particles to the edge as evaporation proceeds. After evaporation, suspended particles are left highly concentrated along the original drop edge. The coffee-ring effect is manifested in systems with diverse constituents, ranging from large colloids[1,4,5] to nanoparticles[6] and individual molecules[7]. In fact—despite the many practical applications for uniform coatings in printing[8], biology[9,10] and complex assembly[11]—the ubiquitous nature of the effect has made it difficult to avoid[6,12–16]. Here we show experimentally that the shape of the suspended particles is important and can be used to eliminate the coffee-ring effect: ellipsoidal particles are deposited uniformly during evaporation. The anisotropic shape of the particles significantly deforms interfaces, producing strong interparticle capillary interactions[17–23]. Thus, after the ellipsoids are carried to the air–water interface by the same outward flow that causes the coffee-ring effect for spheres, strong long-ranged interparticle attractions between ellipsoids lead to the formation of loosely packed or arrested structures on the air–water interface[17,18,21,24]. These structures prevent the suspended particles from reaching the drop edge and ensure uniform deposition. Interestingly, under appropriate conditions, suspensions of spheres mixed with a small number of ellipsoids also produce uniform deposition. Thus, particle shape provides a convenient parameter to control the deposition of particles, without modification of particle or solvent chemistry.**

A drop of evaporating water is a complex, difficult-to-control, non-equilibrium system. Along with capillary flow, the evaporating drop features an air–water interface shaped like a spherical cap and also Marangoni flows induced by small temperature differences between the top of the drop and the contact line[4]. Attempts to reverse or ameliorate the coffee-ring effect have thus far focused on manipulating capillary flows[6,12–16]. In this contribution, we show that uniform coatings during drying can be obtained simply by changing particle shape. The uniform deposition of ellipsoids after evaporation (Fig. 1a) is readily apparent, and stands in stark contrast to the uneven 'coffee ring' deposition of spheres (Fig. 1b) in the same solvent, with the same chemical composition, and experiencing the same capillary flows (Fig. 1c).

Much of the physics of the coffee-ring effect has been demonstrated with micrometre-sized polystyrene particles[1]. Here we also utilize such particles and simply modify their shape. Our experiments use water drops containing a suspension of micrometre-sized polystyrene spheres stretched asymmetrically to different aspect ratios[25,26]. We note that similar results were obtained for hydrophilic ellipsoids and other anisotropic particles (see Supplementary Fig. 4). We evaporate the drops on glass slides and study suspensions containing particles of the same composition, but with different major-axis/minor-axis aspect ratios ($\alpha$), including spheres ($\alpha = 1.0$), slightly deformed spheres ($\alpha = 1.05$, 1.1, 1.2, 1.5) and ellipsoids ($\alpha = 2.5$, 3.5); we study volume fractions ($\phi$) that vary from $10^{-4}$ to $2 \times 10^{-1}$.

During the drying process, the droplet contact line remains pinned in all suspensions, and fluid (carrying particles) flows outward from the drop centre to replenish the edges. Spherical particles are efficiently transported to the edge, either in the bulk or along the air–water interface, leaving a ring after evaporation is complete (Supplementary

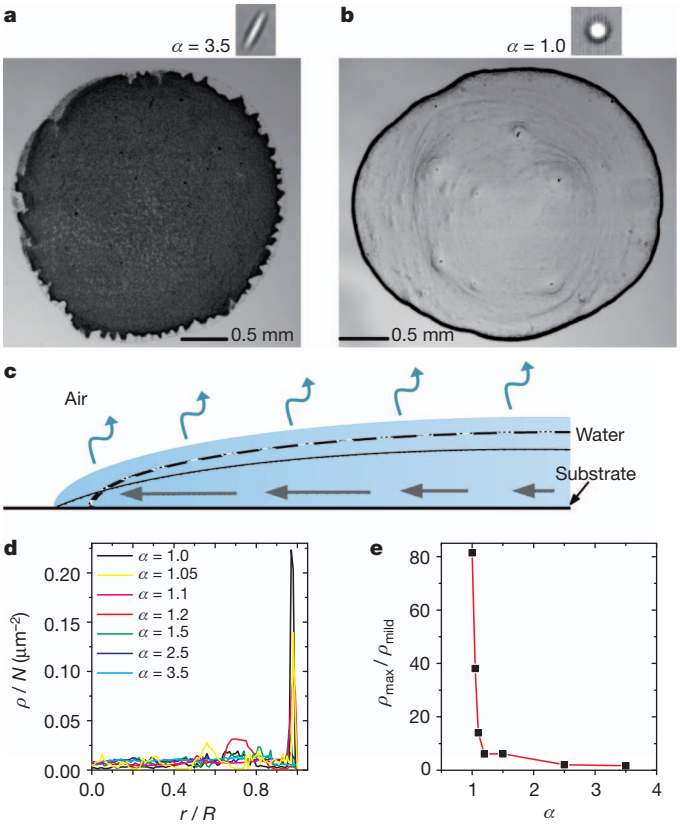



**Figure 1 | Deposition of spheres and ellipsoids.** **a**, **b**, Images of the final distributions of ellipsoids (**a**) and spheres (**b**) after evaporation. Insets, particle shape. **c**, Schematic diagram of the evaporation process, depicting capillary flow induced by pinned edges. Evaporation occurs over the entire drop surface (blue arrows), so if the contact line were free to recede, the drop profile would be preserved during evaporation (dashed line). However, the contact line remains pinned, so the contact angle decreases (solid line). Thus, a capillary flow (black arrows), from the drop's centre to its edges, is induced to replenish fluid at the contact line. **d**, Droplet-normalized particle number density, $\rho/N$, plotted as function of radial distance from centre of drop for ellipsoids with various major–minor axis aspect ratios ($\alpha$). **e**, The maximum local density, $\rho_{max}$, normalized by the density in the middle of the drop, $\rho_{mid}$, plotted for all $\alpha$. Red lines guide the eye.

[1]Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [2]Complex Assemblies of Soft Matter, CNRS-Rhodia-University of Pennsylvania UMI 3254, Bristol, Pennsylvania 19007, USA.



©2011 Macmillan Publishers Limited. All rights reserved

Movies 1, 3 and 5). Anisotropic particles ($\alpha > 1.0$), however, are only transported towards the edge until they reach the air–water interface (Supplementary Movies 2, 4 and 6). Once at that interface, ellipsoids experience strong long-ranged attractions to other ellipsoids[17–22], leading to the formation of loosely packed quasi-static or arrested structures at the interface[24]. On the interface, the interparticle attraction between anisotropic particles is more than two orders of magnitude stronger than the attraction between spheres[23]. Thus, anisotropic particles in these 'open' structures are strongly bound to each other and to the interface, so the energy cost of deforming, moving or breaking up these clusters is very large. As a result, ellipsoid mobility is markedly reduced, and they resist the radially outward flow (Supplementary Information). Thus, when evaporation is complete, anisotropic particles are much more uniformly deposited on the glass surface than spheres. Although spheres also adsorb onto the interface during evaporation, they do not significantly deform the interface[17]. Therefore, the radially outward fluid flow continues to push them to the drop's edge[4]. Figure 1 shows the final deposition, after evaporation on glass slides at 23 °C, of suspensions of particles ($\phi = 0.005$, 1 μl) with different aspect ratios. Spherical particles are primarily deposited at the original perimeter of the droplet (Fig. 1b). Ellipsoidal particles are distributed much more uniformly (Fig. 1a).

To quantify the behaviour shown qualitatively in Fig. 1a and b, we determined the areal number fraction, $\rho(r)$, of particles deposited as a function of radial distance ($r$) from the drop centre, normalized by the number of particles, $N$ (Fig. 1d). For spheres ($\alpha = 1.0$), $\rho/N$ is ~70 times larger at $r/R \approx 1$ than in the middle of the drop (here $R$ is the drop radius). Conversely, the density profile of ellipsoidal particles is fairly uniform as a function of $r/R$. As aspect ratio is increased from unity, the peak at large $r/R$ decreases. A clear coffee-ring effect persists for particles only marginally distorted from their original spherical shape ($\alpha = 1.05$ and 1.1), but particles slightly more anisotropic ($\alpha = 1.2$) are deposited uniformly. This behaviour is summarized by calculating $\rho_{max}/\rho_{mid}$ (Fig. 1e), where $\rho_{max}$ is the maximum value of $\rho$ (typically located at $r/R \approx 1$) and $\rho_{mid}$ is the average value of $\rho$ in the middle of the drop ($r/R < 0.25$). (See Supplementary Information and Supplementary Fig. 6.)

Quantification of the spatio-temporal evaporation profile of the suspensions provides a first step towards understanding why ellipsoids are deposited uniformly. We measure drop mass of different suspensions during evaporation (Supplementary Fig. 1). The drop mass decreases linearly in time, and the mass rate-of-change (10.0 μg s$^{-1}$) is the same for drops of sphere suspension, drops of ellipsoid suspension, and drops of pure water[1,5].

To confirm that the contact line remains pinned until the final stage of evaporation, we used video microscopy to measure drop radius during evaporation (Supplementary Fig. 2). Using these data, the time at which evaporation finishes, $t_{final}$, is readily identified as the time when the drop radius shrinks to zero. Interestingly, we find that the radius decreases by less than 10% until $t = 0.8t_{final}$; that is, the contact line is pinned for the vast majority of the evaporation time period in all samples. For suspensions of ellipsoids, the contact line becomes partially depinned around $t = 0.7t_{final}$, but does not become completely depinned until $t = 0.8t_{final}$. The experiments thus demonstrate that despite similar contact line behaviour, capillary flow, and evaporation rates, the deposition of spheres and ellipsoids differs significantly.

Experimental images clearly reveal that whereas spheres are carried to the drop's edge (Fig. 2a–d), ellipsoids are carried there to far lesser degree (Fig. 2e–h) (Supplementary Movies 2 and 4). We measured the average areal particle density close to the contact line, $\rho_R = \int_{r=R-20\mu m}^{r=R} \rho(r)dr$, as a function of time (Fig. 2i), and thereby demonstrated that ellipsoid density grows at a slower rate than sphere density (Supplementary Information.)

Images of particles near the drop's contact line (Fig. 2j–m) reveal that unlike spheres, which are carried from the bulk all the way to the contact line, most ellipsoids adhere to the loosely packed structures at

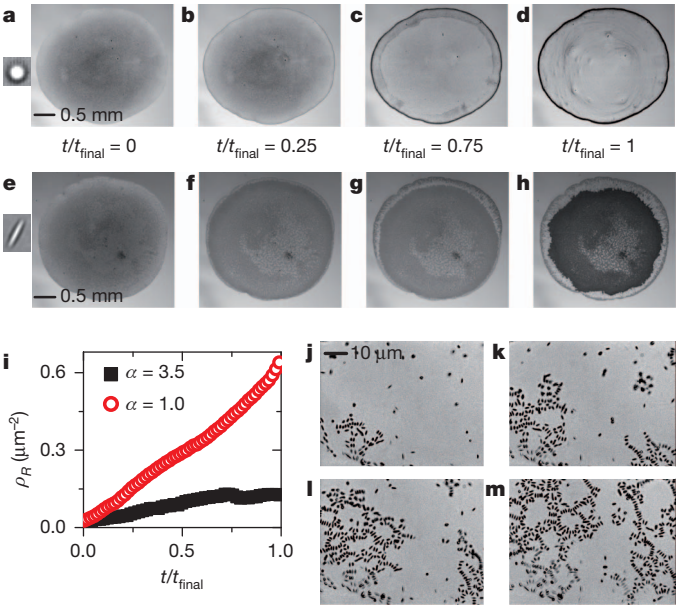



**Figure 2 | Transportation of particles over time.** **a–h**, Experimental snapshots at different times $t/t_{final}$ during the evaporation of a drop of particle suspension; shown are data for spherical particles (left inset: **a–d**) and for ellipsoidal particles with aspect ratio $\alpha = 3.5$ (left inset: **e–h**). **i**, The areal particle density, $\rho_R$, located within 20 μm of the contact line (that is, the drop edge) as a function of time during evaporation for ellipsoidal particles. **j–m**, Images of the assembly of ellipsoids at the air–water interface over the same time intervals during evaporation. Loosely packed structures form on the air–water interface, preventing ellipsoids from reaching the edge. The three-phase contact line is at the bottom left corner of these snapshots.

the air–water interface before they reach the contact line (see below). This capillary attraction has been characterized in prior experiments as long-ranged and very strong[17,18,23,27–29]. The loosely packed configurations formed by ellipsoids on the interface are structurally similar to those seen in previous experiments on ellipsoids at flat air–water and water–oil interfaces[17,18,21]. They produce a surface viscosity that is much larger than the bulk viscosity, facilitating ellipsoid resistance to radially outward flows (Supplementary Information). We note that spheres also adsorb onto the interface during evaporation. However, spheres do not strongly deform the interface[17] and they experience a much weaker interparticle attraction than ellipsoids[23]; thus, radially outward fluid flows push spheres to the drop's edge[4].

Snapshots of the region within 40 μm of the drop contact line confirm that whereas spheres pack closely at the edge (Fig. 3a), ellipsoids form loosely packed structures (Fig. 3b), which prevent particles from reaching the contact line (Supplementary Movies 5 and 6). Particles with $\alpha = 1.2$ and 1.5 pack at higher area fractions than ellipsoids with $\alpha > 1.5$, resulting in larger values of $\rho_{max}/\rho_{mid}$ for $\alpha = 1.2$ and 1.5 and producing the small peak in $\rho(r)$ at $r/R = 0.7$ for $\alpha = 1.2$. The structures on the air–water interface appear to be locally arrested or 'jammed'[24], that is, particles do not rearrange. Once an ellipsoid joins a collective structure, its position relative to other ellipsoids changes by less than 20 nm (the lower limit of our resolution), and the structure only rearranges when new particles become attached to the interface. In order to quantify the ability of interfacial aggregates of ellipsoids to resist bulk flow, we calculated the Boussinesq number for our experimental system (see Supplementary Information for details), that is, the ratio of the surface drag to bulk drag for ellipsoids with $\alpha = 3.5$. The calculated Boussinesq number is much greater than 1, and grows exponentially with time (Supplementary Fig. 3). This behaviour is expected, as shear stress grows linearly with particle velocity, but elastic modulus grows exponentially with ellipsoid area fraction[21], and dominates the ratio. As might be anticipated, the Boussinesq number for spheres was much smaller than for ellipsoids, that is, it is less than 1.



©2011 Macmillan Publishers Limited. All rights reserved

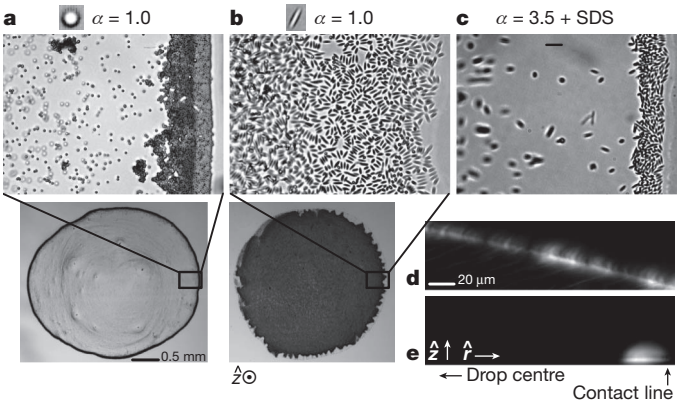



**Figure 3 | High-magnification images of particles near the drop contact line. a–c**, Top, microscope images of a region within 40 μm of the drop contact line, taken at time $t/t_{final} = 0.5$, for suspensions of spheres (**a**), ellipsoids ($\alpha = 3.5$) (**b**), and ellipsoids ($\alpha = 3.5$) mixed with surfactant (SDS; 0.2 wt%) (**c**). Spheres pack closely at the contact line. Ellipsoids form loosely packed structures. Surfactant lowers the drop surface tension, making ellipsoids pack closely at the contact line. For **a** and **b**, pictures of the entire drop after evaporation are shown (bottom) and the magnified region is boxed. **d**, **e**, Confocal projections of suspensions of ellipsoids ($\alpha = 2.5$) (**d**) and spheres (**e**) onto the $z$–$r$ plane in cylindrical coordinates. This $z$-direction is indicated below **b**.

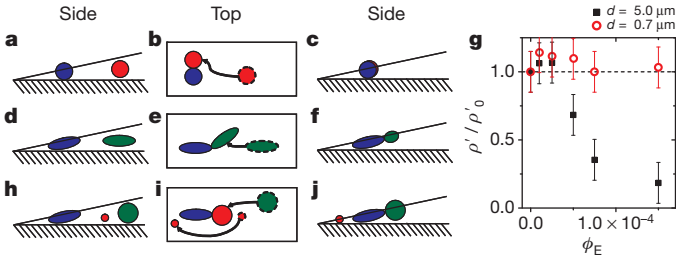



**Figure 4 | Behaviour of spheres, ellipsoids, and mixtures of spheres and ellipsoids in drying liquid drops.** For all cartoons, the leftmost and rightmost panels are side views at early and late times, respectively, and the centre panel is a top view showing particle trajectories (arrows) linking those times (original positions indicated with dashed lines, final positions indicated with solid lines). **a–f**, Cartoons depicting capillary flow that carries spheres (**a–c**) and ellipsoids (**d–f**) to the drop's edge. Spheres leave a ring-like formation, whereas ellipsoids form loosely packed structures on the air–water interface. **g**, The deposition of mixtures of spheres and ellipsoids is characterized by the ratio $\rho' = \rho_{max}/\rho_{mid}$ (where $\rho_{max}$ is the maximum local density and $\rho_{mid}$ is the density in the middle of the drop) as a function of ellipsoid volume fraction, $\phi_E$. Two sizes of particles are studied: $d = 5.0$ μm (black squares) and $d = 0.7$ μm (red circles), where $d$ is the particle diameter. $\rho'$ is normalized by $\rho'_0$, the value of $\rho'$ when there are no ellipsoids present, that is, when $\phi_E = 0$. The dashed line represents no change in the distribution of particles. The coffee-ring effect persists for mixtures of small spheres and ellipsoids, but is destroyed for mixtures of large spheres and ellipsoids. Error bars represent the uncertainty that results from finite bin sizes. **h–j**, Cartoons depicting capillary flow that carries suspensions of spheres and ellipsoids to the drop's edge.

To confirm that deformations of the interface are responsible for the uniform deposition of ellipsoids, we added a small amount of surfactant (sodium dodecyl sulphate, SDS, 0.2% by weight) to a suspension of ellipsoids with $\alpha = 3.5$. Surfactant lowers the surface tension of the drop, thus making interfacial deformations less energetically costly and shorter-range. This restores the coffee-ring effect; ellipsoids pack closely at the contact line (Fig. 3c), in a manner similar to spheres (Supplementary Movie 7). The ellipsoids no longer strongly deform the air–water interface and their interactions with other ellipsoids are correspondingly reduced; as a result, they move more easily along and on-and-off the interface and are able to pack close to the contact line. (We note that, without added surfactant, ellipsoids increase the air–water surface tension, as evidenced by an increase in contact angle from ~15° for spheres to ~35° for ellipsoids with $\alpha = 3.5$; Supplementary Fig. 5).

We obtained direct evidence that the ellipsoids sit at the air–water interface, using three-dimensional confocal microscopy. Confocal snapshots, projected onto a side view of the drop, clearly confirm that ellipsoids sit at the air–water interface (Fig. 3d), whereas spheres do not and are carried all the way to the contact line (Fig. 3e, summarized in Fig. 4a–f).

Finally, we showed that the addition of small numbers of ellipsoids to sphere suspensions can also destroy the coffee-ring effect. If the diameter of the spheres is smaller than the minor axis of the ellipsoids, then the coffee ring persists; if the diameter is larger than that axis, then the coffee ring is destroyed. To observe this effect, we evaporated drops of suspensions containing both ellipsoids and spheres.

Each suspension contains ellipsoids with $\alpha = 3.5$, stretched from particles of diameter $d = 1.3$ μm, and spheres suspended at a volume fraction $\phi = 0.02$. Evaporative deposits are characterized as a function of ellipsoid volume fraction $\phi_E$ via $\rho'(\phi_E) = \rho_{max}/\rho_{mid}$ (Fig. 4g). First, we evaporated suspensions containing smaller spheres with $d = 0.7$ μm, along with the ellipsoids at volume fractions ranging from $\phi_E = 0$ to $1.5 \times 10^{-4}$. After evaporation, the spheres displayed a clear coffee ring, and this coffee ring persists even if more ellipsoids are added to the initial suspension (Fig. 4g). The coffee-ring effect is uninhibited because spheres that are smaller than the ellipsoids are easily able to move under or through the loosely packed particle structures and reach the drop's edge (Fig. 4h–j).

If, instead, we evaporate suspensions containing larger spheres with $d = 5.0$ μm, along with the same ellipsoids at the same volume fractions used previously, then different phenomena emerge. When the ellipsoid volume fraction is very small ($\phi_E \leq 2.5 \times 10^{-5}$), the suspensions still exhibit a clear coffee-ring effect. However, at larger $\phi_E$, the coffee ring is diminished, and it eventually disappears at sufficiently large $\phi_E$, that is, $\phi_E \approx 1.5 \times 10^{-4}$ (Fig. 4g). In this case, the larger spheres adsorb onto the air–water interface farther from the drop edge than do the ellipsoids. In the absence of ellipsoids, the spherical particles form closely packed aggregates, but in the presence of ellipsoids they instead join the loosely packed aggregates, eliminating the coffee-ring effect (Fig. 4h–j). Thus, uniform depositions can potentially be made with existing suspensions, simply by adding ellipsoids.

Looking forward, we note that the ability to deposit particles uniformly is desirable in many applications[8]. Unfortunately, most proposed methods for avoiding the coffee-ring effect require long multistage processes, which can be costly in manufacturing or require use of organic solvents which are sometimes flammable and toxic[6,12]. Here we have shown that by exploiting a particle's shape, a uniform deposit can be easily derived from an evaporating aqueous solution. These results further suggest that other methods of inducing strong capillary interactions, for example, surface roughness[30], may also produce uniform deposits.

## METHODS SUMMARY

To create ellipsoidal particles, 1.3-μm-diameter polystyrene particles (Invitrogen) are suspended in a polyvinyl alcohol (PVA) gel and are heated above the polystyrene melting point (~100 °C), but below the PVA melting point (~180 °C)[25,26]. Polystyrene melts in the process, but the PVA gel only softens. The PVA gel is then pulled so that the spherical cavities containing liquid polystyrene are stretched into ellipsoidal cavities. When the PVA gel cools, polystyrene solidifies in the distorted cavities and becomes frozen into an ellipsoidal shape. The hardened gel is dissolved in water, and the PVA is removed via centrifugation (see Supplementary Information for more details about PVA removal). Each iteration of this process creates ~$10^9$ ellipsoidal particles in ~50-μl suspensions. The particles are charge-stabilized, and the resultant suspensions are surfactant-free. Snapshots of experimental particles are shown in Fig. 1a and b insets. The aspect ratio polydispersity is ~10%. To ensure the preparation process does not affect particle deposition, our spheres undergo the same procedure, without stretching. In our experiments, suspensions of particles are placed on glass slides (Fisher Scientific) and evaporated.

Confocal snapshots are shown in Fig. 3d and e. By integrating the brightness of each pixel over a period of 0.05 s, only particles that are roughly stationary during



©2011 Macmillan Publishers Limited. All rights reserved

this time period appear in the images. Snapshots are then projected onto a side-view of the drop.

**Received 12 April; accepted 4 July 2011.**


1. Deegan, R. D. *et al.* Capillary flow as the cause of ring stains from dried liquid drops. *Nature* **389,** 827–829 (1997).
2. Denkov, N. D. *et al.* Two-dimensional crystallization. *Nature* **361,** 26 (1993).
3. Hu, H. & Larson, R. G. Evaporation of a sessile droplet on a substrate. *J. Phys. Chem. B* **106,** 1334–1344 (2002).
4. Deegan, R. D. *et al.* Contact line deposits in an evaporating drop. *Phys. Rev. E* **62,** 756–765 (2000).
5. Deegan, R. D. Pattern formation in drying drops. *Phys. Rev. E* **61,** 475–485 (2000).
6. Bigioni, T. P. *et al.* Kinetically driven self assembly of highly ordered nanoparticle monolayers. *Nature Mater.* **5,** 265–270 (2006).
7. Kaya, D., Belyi, V. A. & Muthukumar, M. Pattern formation in drying droplets of polyelectrolyte and salt. *J. Chem. Phys.* **133,** 114905 (2010).
8. Park, J. & Moon, J. Control of colloidal particle deposit patterns within picoliter droplets ejected by ink-jet printing. *Langmuir* **22,** 3506–3513 (2006).
9. Dugas, V. Immobilization of single-stranded DNA fragments to solid surfaces and their repeatable specific hybridization: covalent binding or adsorption? *Sens. Actuators B* **101,** 112–121 (2004).
10. Dugas, V., Broutin, J. & Souteyrand, E. Droplet evaporation study applied to DNA chip manufacturing. *Langmuir* **21,** 9130–9136 (2005).
11. de Gans, B. J., Duineveld, P. C., & Schubert, U. S. Inkjet printing of polymers: state of the art and future developments. *Adv. Mater.* **16,** 203–213 (2004).
12. Hu, H. & Larson, R. G. Marangoni effect reverses coffee-ring depositions. *J. Phys. Chem. B* **110,** 7090–7094 (2006).
13. Weon, B. M. & Je, J. H. Capillary force repels coffee-ring effect. *Phys. Rev. E* **82,** 015305 (2010).
14. Kajiya, T., Kobayashi, W., Okuzono, T. & Doi, M. Controlling the drying and film formation processes of polymer solution droplets with addition of small amount of surfactants. *J. Phys. Chem. B* **113,** 15460–15466 (2009).
15. Nguyen, V. X. & Stebe, K. J. Patterning of small particles by a surfactant-enhanced Marangoni-Benard instability. *Phys. Rev. Lett.* **88,** 164501 (2002).
16. Park, B. J. & Furst, E. M. Fluid-interface templating of two-dimensional colloidal crystals. *Soft Matter* **6,** 485–488 (2010).
17. Loudet, J. C., Yodh, A. G. & Pouligny, B. Wetting and contact lines of micrometer-sized ellipsoids. *Phys. Rev. Lett.* **97,** 018304 (2006).
18. Loudet, J. C., Alsayed, A. M., Zhang, J. & Yodh, A. G. Capillary interactions between anisotropic colloidal particles. *Phys. Rev. Lett.* **94,** 018301 (2005).
19. Bowden, N., Arias, F., Deng, T. & Whitesides, G. M. Self-assembly of microscale objects at a liquid/liquid interface through lateral capillary forces. *Langmuir* **17,** 1757–1765 (2001).
20. Brown, A. B. D., Smith, C. G. & Rennie, A. R. Fabricating colloidal particles with photolithography and their interactions at an air-water interface. *Phys. Rev. E* **62,** 951–960 (2000).
21. Madivala, B., Fransaer, J. & Vermant, J. Self-assembly and rheology of ellipsoidal particles at interfaces. *Langmuir* **25,** 2718–2728 (2009).
22. Madivala, B., Vandebril, S., Fransaer, J. & Vermant, J. Exploiting particle shape in solid stabilized emulsions. *Soft Matter* **5,** 1717–1727 (2009).
23. Park, B. J. & Furst, E. M. Attractive interactions between colloids at the oil-water interface. *Soft Matter* http://dx.doi.org/10.1039/c1sm00005e (published online, 26 April 2011).
24. Fournier, J. B. & Galatola, P. Anisotropic capillary interactions and jamming of colloidal particles trapped at a liquid-fluid interface. *Phys. Rev. E* **65,** 031601 (2002).
25. Champion, J. A., Katare, Y. K. & Mitragotri, S. Making polymeric micro- and nanoparticles of complex shapes. *Proc. Natl Acad. Sci. USA* **104,** 11901–11904 (2007).
26. Ho, C. C., Keller, A., Odell, J. A. & Ottewill, R. H. Preparation of monodisperse ellipsoidal polystyrene particles. *Colloid Polym. Sci.* **271,** 469–479 (1993).
27. Lehle, H., Noruzifar, E. & Oettel, M. Ellipsoidal particles at fluid interfaces. *Eur. Phys. J. E* **26,** 151–160 (2008).
28. Lewandowski, E. P., Bernate, J. A., Tseng, A., Searson, P. C. & Stebe, K. J. Oriented assembly of anisotropic particles by capillary interactions. *Soft Matter* **5,** 886–890 (2009).
29. Danov, K. D. & Kralchevsky, P. A. Capillary forces between particles at a liquid interface: general theoretical approach and interactions between capillary multipoles. *Adv. Colloid Interface Sci.* **154,** 91–103 (2010).
30. Stamou, D., Duschl, C. & Johannsmann, D. Long-range attraction between colloidal spheres at the air-water interface: the consequence of an irregular meniscus. *Phys. Rev. E* **62,** 5263–5272 (2000).


**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank K. B. Aptowicz, A. Basu, E. Buckley, J. C. Crocker, P. Habdas, T. Lubensky and K. J. Stebe for discussions and critical reading of the manuscript. We thank J. Crassous and H. Dietsch for the hydrophilic ellipsoids and spheres (see Supplementary Fig. 4). We acknowledge financial support from the National Science Foundation through DMR-0804881, the PENN MRSEC DMR-0520020 and NASA NNX08A00G. T.S. thanks DAAD for a postdoctoral fellowship.

**Author Contributions** P.J.Y. initiated the research, designed and performed the experiments, synthesized the colloidal particles, and analysed the data with support from A.G.Y., T.S. and M.A.L.; P.J.Y. and A.G.Y. wrote the manuscript, with support from T.S. and M.A.L.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. Readers are welcome to comment on the online version of this article at www.nature.com/nature. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.J.Y. (pyunker@sas.upenn.edu).



©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10310

# Southern Ocean dust–climate coupling over the past four million years

Alfredo Martínez-Garcia[1,2,3], Antoni Rosell-Melé[3,4,5], Samuel L. Jaccard[1], Walter Geibert[6,7], Daniel M. Sigman[8] & Gerald H. Haug[1,2]

**Dust has the potential to modify global climate by influencing the radiative balance of the atmosphere and by supplying iron and other essential limiting micronutrients to the ocean[1,2]. Indeed, dust supply to the Southern Ocean increases during ice ages, and 'iron fertilization' of the subantarctic zone may have contributed up to 40 parts per million by volume (p.p.m.v.) of the decrease (80–100 p.p.m.v.) in atmospheric carbon dioxide observed during late Pleistocene glacial cycles[3–7]. So far, however, the magnitude of Southern Ocean dust deposition in earlier times and its role in the development and evolution of Pleistocene glacial cycles have remained unclear. Here we report a high-resolution record of dust and iron supply to the Southern Ocean over the past four million years, derived from the analysis of marine sediments from ODP Site 1090, located in the Atlantic sector of the subantarctic zone. The close correspondence of our dust and iron deposition records with Antarctic ice core reconstructions of dust flux covering the past 800,000 years (refs 8, 9) indicates that both of these archives record large-scale deposition changes that should apply to most of the Southern Ocean, validating previous interpretations of the ice core data. The extension of the record beyond the interval covered by the Antarctic ice cores reveals that, in contrast to the relatively gradual intensification of glacial cycles over the past three million years, Southern Ocean dust and iron flux rose sharply at the Mid-Pleistocene climatic transition around 1.25 million years ago. This finding complements previous observations over late Pleistocene glacial cycles[5,8,9], providing new evidence of a tight connection between high dust input to the Southern Ocean and the emergence of the deep glaciations that characterize the past one million years of Earth history.**

The scarcity of iron reduces marine export production and carbon uptake in one-quarter of the world ocean where the concentration of major nutrients (phosphorus and nitrogen) is perennially high[10] (Fig. 1). The Southern Ocean is the region where variations in iron availability can have the largest effect on Earth's carbon cycle through its fertilizing effect on marine ecosystems.

In the modern Southern Ocean, deep ocean waters with high $CO_2$ and nutrient content are brought to the surface by wind-driven upwelling and density-driven overturning. However, the scarcity of iron reduces phytoplankton growth[1,10], and major nutrients are returned to the subsurface before they are fully consumed. This incomplete utilization of nutrients represents an important leak in the modern global biological pump because it allows the escape of deeply sequestered carbon back to the atmosphere, thereby raising atmospheric $CO_2$ levels[6]. Today, the Antarctic zone is the greater part of the Southern Ocean $CO_2$ leak, because of its extremely low degree of surface nutrient consumption and its ventilation of the abyssal ocean. The subantarctic zone, further north in the Southern Ocean, is characterized by greater nutrient consumption and ventilates the somewhat smaller volume of the mid-depth ocean. However, its nutrient status affects the productivity of the low latitude ocean, with repercussions for low-latitude $CaCO_3$ production and for the partitioning of regenerated carbon storage between the mid-depth and deep ocean. Thus, both the Antarctic zone and the subantarctic zone are potentially important for glacial/interglacial $CO_2$ change[6,7].

We first focus on the Antarctic zone. An increase in the efficiency of the global biological pump can be accomplished either by decreasing the physical cycling of deep water through the surface of the Antarctic zone[11] or by increasing the degree to which Antarctic zone surface nutrients are consumed by marine organisms (for example, as a result of an increase in the availability of iron relative to other nutrients[6,12]). Data and models have been used to argue that the Antarctic was more strongly stratified during ice ages, reducing the Antarctic leak in the biological pump[6,12]. Because Antarctic major nutrient and iron supply come from upwelled water, productivity in the Antarctic zone would probably decline during ice ages in step with stratification, consistent with most Antarctic productivity data[4]. Nevertheless, the greater the input of aeolian iron, the greater the degree of nutrient consumption is expected for a given increase in Antarctic stratification, because the dust input raises the total iron supply to the surface ocean relative to the input of major nutrients and excess $CO_2$ from below[6,7]. Thus, dust input to the Antarctic is a central regulator of how completely the Antarctic $CO_2$ leak can be sealed during ice ages, regardless of whether productivity increased or decreased in the region.

Second, we consider the subantarctic zone. This zone, unlike the Antarctic, is characterized by higher productivity during ice ages[4,5,13].

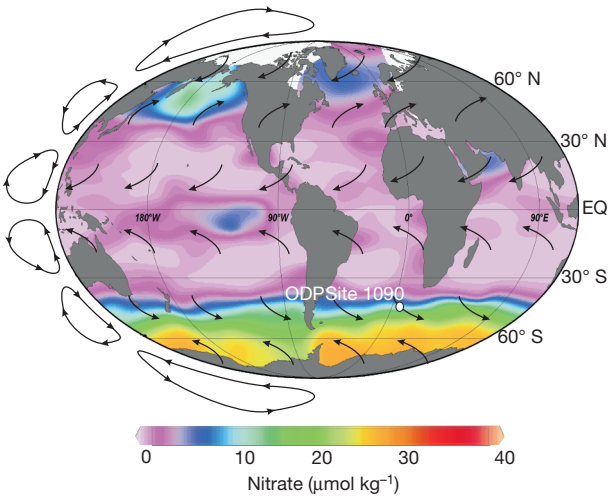



**Figure 1 | Location of ODP Site 1090, world ocean surface nitrate concentrations, and wind direction.** ODP Site 1090, the source of the sediment cores used in this study, is located at 42° 54.5′ S, 8° 54.0′ E, at 3,702 m depth. Nitrate concentrations are from the Electronic Atlas of the World Ocean Experiment. Black arrows are schematic representations of atmospheric convection cells and wind directions. EQ, Equator.

[1]Geological Institute, ETH Zürich, Zürich 8092, Switzerland. [2]DFG-Leibniz Center for Surface Process and Climate Studies, Institute for Geosciences, Potsdam University, Potsdam D-14476, Germany. [3]Institut de Ciència i Tecnologia Ambientals (ICTA), Universitat Autònoma de Barcelona, Bellaterra 08193, Catalonia, Spain. [4]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona 08010, Catalonia, Spain. [5]College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331-5503, USA. [6]School of GeoSciences, The University of Edinburgh, Edinburgh EH9 3JW, UK. [7]Scottish Association for Marine Science (SAMS), Scottish Marine Laboratory, Oban, Argyll PA37 1QA, UK. [8]Department of Geosciences, Princeton University, New Jersey 08544, USA.



©2011 Macmillan Publishers Limited. All rights reserved

This suggests a conceptually simpler mechanism than in the Antarctic zone by which dust-driven iron fertilization of the subantarctic zone may have lowered atmospheric $CO_2$, with the increase in subantarctic zone productivity consuming a greater fraction of the surface nutrients and thus driving more storage of carbon in the ocean interior. The subantarctic zone is one of the best candidates for aeolian iron fertilization, because it falls in the same latitude as the major Southern Hemisphere dust sources, and is on the wind path from them[3,6]. Further, there is evidence that natural aeolian iron deposition enhances marine export production in the modern subantarctic zone[14]. Palaeoceanographic data appear to constrain its effect to a maximum of 40 p.p.m.v. of the atmospheric $CO_2$ decrease observed in the last glacial cycle, because a significant increase in dust fluxes and marine export production in the subantarctic zone is only observed in the transition from marine isotope stage (MIS) 5 to MIS 4, when atmospheric $CO_2$ concentrations are already around 40–50 p.p.m.v. lower than during the interglacial periods[3–7]. A $CO_2$ drawdown of $\leq$40 p.p.m. from subantarctic zone iron fertilization is consistent with estimates obtained using geochemical box models and Earth system models, which in most cases range between 20 and 40 p.p.m.v. (see refs 3, 7, 15–18 and Supplementary Information).

The ice core reconstruction generated by the EPICA project[8,9] has provided a unique record with which to test dust–climate interactions in the Southern Ocean over the past 800,000 years (Fig. 2d). However, ice core records only allow an indirect inference of the supply of dust and iron to the ocean, and it is unclear how far back in time they can be extended[19]. Here we report a high-resolution dust and iron record from the subantarctic Atlantic spanning the past four million years (4 Myr), which provides the first insights into the variability of Southern Ocean dust deposition through the major climatic transitions of the Pliocene and Pleistocene epochs.

The record is based on the combined analysis of independent organic (*n*-alkanes) and inorganic (Fe, Ti and $^{232}$Th) dust proxies in marine sediments from ODP Site 1090 (Fig. 1). Long-chain *n*-alkanes are lipid constituents of the epicuticular wax layer of terrestrial plants. These terrestrial lipids are eroded from leaf surfaces and soils by winds and transported in the organic fraction of aeolian dust over the ocean[20]. In our record, there is a clear predominance of long-chain *n*-alkanes with odd carbon numbers over those with even carbon numbers, over the Pliocene and Pleistocene epochs. This is representative of inputs of leaf waxes from terrestrial plants transported by wind[20], supporting the interpretation of the record in terms of changes in dust deposition through time[5]. Fe, Ti and $^{232}$Th are elements with a well-characterized lithogenic origin. Hence, their concentration in marine sediments can be used to quantitatively estimate variations in the supply of wind-borne lithogenic material to remote areas of the ocean where other sources of continental material are negligible[5,13,21] (see Supplementary Information). The consistency among the different dust tracers measured in ODP Site 1090 and their excellent correlation with the dust deposition record from the Antarctic ice cores over the past 0.8 Myr (Fig. 2d) indicate that they are faithful recorders of changes in the supply of dust, and hence of iron, to high southern latitudes through time. This subantarctic zone/ice-core agreement in dust flux history also validates the long-held assumption that the ice core records reflect the pattern of dust fluxes to the Southern Ocean surface. A difference in the magnitude of the glacial–interglacial change of dust flux between ocean sediment and polar ice core records has been noted previously[5,8]. We note a similar difference here, with our record indicating a smaller glacial–interglacial change compared with the changes of more than an order of magnitude indicated by the Antarctic ice cores (Fig. 2d). This difference is probably due to variations in the efficiency of dust transport mechanisms to the Antarctic continent during glacial stages[8].

Our record reveals that, on orbital timescales, Southern Ocean dust deposition has been exponentially linked to global ice volume changes through the Pliocene and Pleistocene epochs (Fig. 2). Major steps in

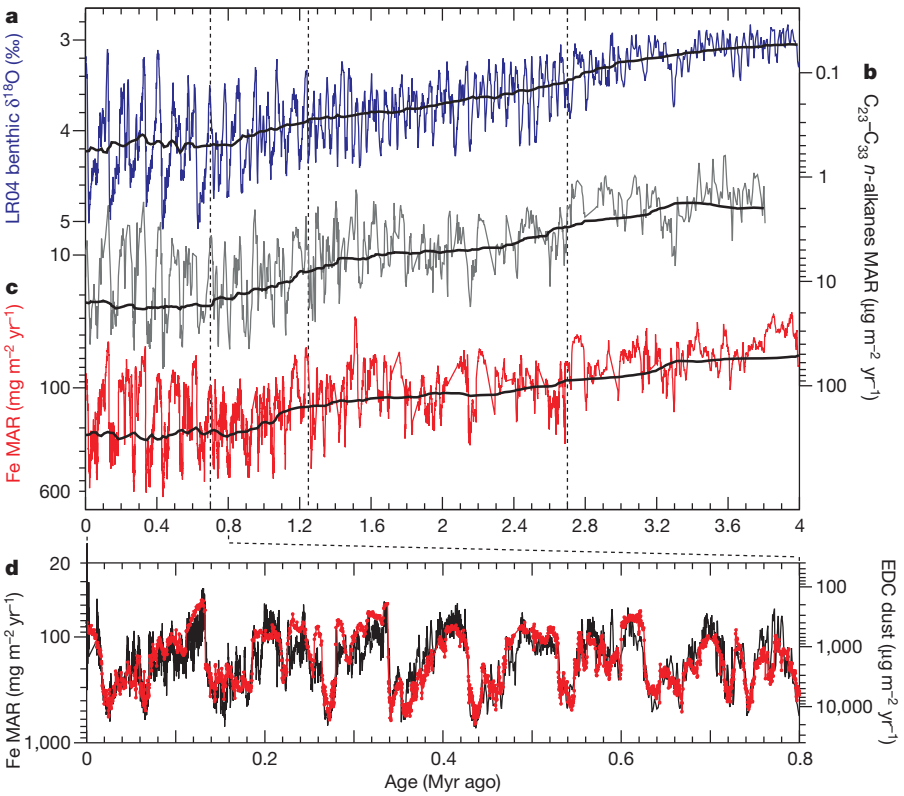



**Figure 2 | Global ice volume, dust and iron deposition in the Southern Ocean since 4 Myr ago.** **a**, Lisiecki and Raymo (LR04) benthic $\delta^{18}$O stack[28]. **b**, **c**, Mass accumulation rates (MARs) of *n*-alkanes (**b**) and iron (**c**) at ODP Site 1090. Both records are plotted on a reversed logarithmic scale to emphasize the exponential relationship with the benthic $\delta^{18}$O record. **d**, Expanded view of iron variability over the past 800 kyr at ODP Site 1090 (red line), plotted together with dust flux estimates in the Antarctic ice core from EPICA Dome C (EDC; black line)[8].



©2011 Macmillan Publishers Limited. All rights reserved

dust deposition coincide with the two major climatic transitions of the Pliocene–Pleistocene, indicating a pervasive link between Southern Ocean dust and Northern Hemisphere glaciations.

The first significant rise in ice-age dust and iron deposition is observed around 2.7 Myr ago, coinciding with the initiation of major glaciations in the Northern Hemisphere (Figs 1 and 2). The development of large, perennial Northern Hemisphere ice sheets around 2.7 Myr ago represents a milestone in climate evolution that marks the end of the relatively stable Pliocene warm period and the emergence of strong feedbacks in the climate system that characterize Quaternary climate oscillations. Although Earth's orbital configuration favoured the initiation of regional glaciations, a decrease in atmospheric $CO_2$ concentrations may be required to initiate and sustain larger glaciations[22]. Data from the Antarctic zone and the subarctic North Pacific suggest that a transition towards permanent polar ocean water-column stratification identified around 2.7 Myr ago may have been a key driver of this $CO_2$ decline[23]. Our data indicate that Antarctic zone stratification may have been reinforced by an increase in iron supply to the Southern Ocean during glacial stages after 2.7 Myr ago, and a similar dynamic may have applied to the subarctic North Pacific.

During the Pliocene, meridional temperature gradients were relatively weak, and the tropical warm pool was apparently more expansive than today[24,25]. This climate state resulted in relatively weak atmospheric convective cells that extended further poleward, greatly affecting wind intensity and global precipitation patterns[24]. As a result of these conditions, models predict stronger precipitation over Africa, Australia and South America during Pliocene times[24], which, together with a weak atmospheric circulation, would have minimized dust generation and transport in these regions. Thus, the strengthening of the meridional temperature gradients at the 2.7-Myr transition and the mid-Pleistocene transition (MPT) as a consequence of the increase in ice volume would have resulted in intensification and spatial contraction of the atmospheric convective cells[24,25] and a reduction in precipitation; these effects would have favoured dust generation and transport, consistent with the increases in dust and iron flux observed at Site 1090.

The most dramatic rise in dust deposition in our record occurs across the MPT (Fig. 3). During this period, Southern Ocean dust fluxes doubled, reaching peak fluxes comparable to those of the Last Glacial Maximum. Although the controls on Southern Ocean dust deposition are not well understood, previous studies have suggested that dust generation requires that a climate threshold must be reached before the system is sensitive to perturbation[5,8,18]. Taking the last glacial cycle as an example, dust flux changed relatively little during the initial climate deterioration during MIS 5, increasing abruptly at MIS 4 with the rapid increase in Northern Hemisphere ice[5,7]. Our long-term dust flux and iron deposition records are strongly convergent with this view, and reveal that Southern Ocean dust flux has a unique sensitivity to the very cold glacial conditions that characterize the past 1.25 Myr. Whereas measures of climate cooling and glaciation indicate a relatively gradual trend in mean climate over the Pleistocene, dust flux at Site 1090 increases exponentially at the MPT, when the most intense ice ages first appear (Fig. 3). The data thus confirm and extend the evidence for a nonlinear relationship of Southern Ocean dust flux to climate (Fig. 3b), indicating that it applies to the entire Pliocene–Pleistocene. This sensitivity of dust flux to climate is also consistent with the observations of relatively constant interglacial tropical sea surface temperature and atmospheric $CO_2$ across the MPT[26,27], as no significant changes in interglacial dust deposition are observed.

Beginning at 1.25 Myr ago, the MPT was characterized by the development of thicker and less stable ice-sheets in the Northern Hemisphere and the emergence of the high-amplitude, asymmetric 100-kyr cycle that distinguishes late Pleistocene glaciations[28,29]. This important climatic shift occurred in the absence of any significant change in orbital forcing, leading to the view that the MPT involved a reorganization of the climate system's internal feedbacks[29]. A

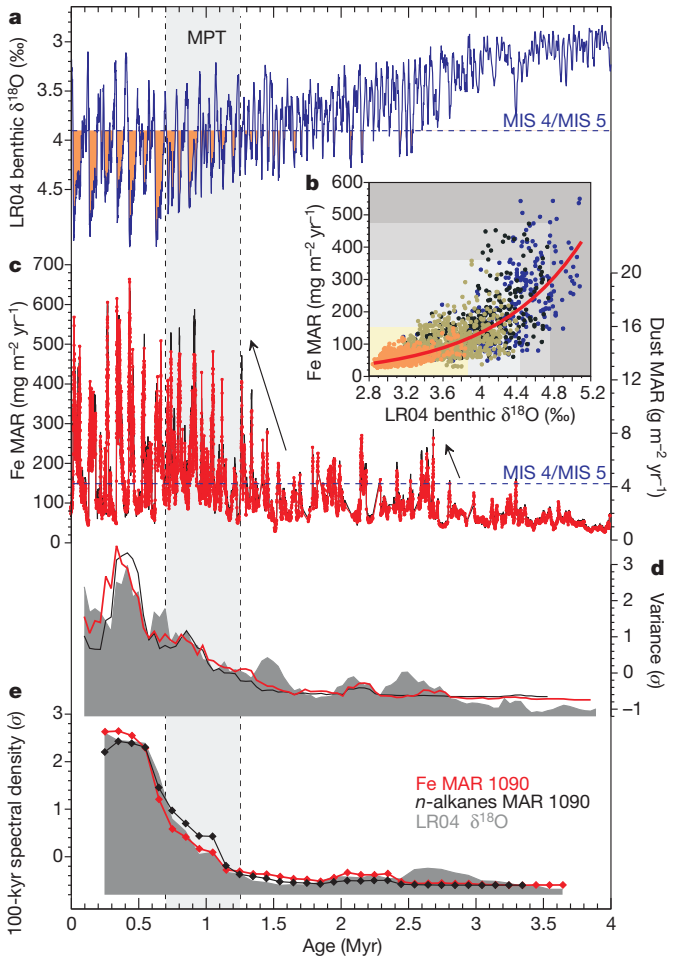



**Figure 3 | Evolution of global ice volume, and Southern Ocean dust and iron variability through the Pliocene and Pleistocene epochs.** **a**, Benthic $\delta^{18}O$ stack[28]. Orange shading indicates the intervals where glaciations intensified and dust fluxes rose to levels comparable to those of the Marine Isotope Stage (MIS) 5 to MIS 4 transition. **b**, Regression between benthic $\delta^{18}O$ and Fe MAR. Colours indicate key time intervals: 0–0.7 Myr ago (blue), 0.7–1.25 Myr ago (black), 1.25–2.7 Myr ago (brown) and 2.7–4 Myr ago (orange). Background shading areas highlight the evolution of the Fe MAR/benthic $\delta^{18}O$ coupling across the defined key time intervals. **c**, Fe MAR (red line), and dust MAR estimated from the Ti MAR (black line). **d**, Normalized variance of Fe MAR, *n*-alkanes MAR, and benthic $\delta^{18}O$, computed in 200-kyr windows with a number of lags equal to half the timeseries length ('½ lags'); key given in **e**. **e**, Normalized power spectral density of the 100-kyr cycle in Fe MAR, *n*-alkanes MAR and $\delta^{18}O$ computed using 500-kyr windows and ½ lags. Vertical dashed lines with grey shading between indicate start and finish of MPT according to ref. 29. In **d** and **e**, the variance and the power spectral density of the 100-kyr cycle of the different records were normalized to zero mean and unit standard deviation ($\sigma$).

decrease in atmospheric $CO_2$ concentrations and the associated global cooling is often proposed as one of the mechanisms to explain the MPT[29]. The available data do indeed suggest that atmospheric $CO_2$ concentrations during peak ice age conditions decreased by ~30 p.p.m. across the MPT[26].

Although our data cannot yet enable us to prove causation, they raise the strong possibility that dust not only increased into the stronger ice ages of the post-MPT world but then also acted as a positive feedback, encouraging further $CO_2$ reduction and cooling through iron fertilization of the Southern Ocean. Consistent with this hypothesis, an increase in opal mass accumulation rate (MAR) has also been observed in the Atlantic sector of the Southern Ocean during this period, coinciding with the development of extensive diatom mats between 47° S and 50° S and with the emergence of the diatom *Fragilariopsis kerguelensis*, the dominant species in modern Southern Ocean sediments[30]. These



©2011 Macmillan Publishers Limited. All rights reserved

observations suggest that an increase in iron availability may have driven part or all of the 30 p.p.m.v. decrease in peak glacial atmospheric $CO_2$ observed across the MPT. By driving the descent into deep, cold glacial periods, the climate/dust/$CO_2$ feedback may underpin the development the strong 100-kyr periodicity that characterizes global climate and dust since the MPT (Fig. 3e). In detail, the greater ice-age dust loads may have increased carbon sequestration by both the Antarctic zone—where it would have encouraged increasingly complete major nutrient consumption even if circulation changes lowered total ice age productivity—and the subantarctic zone, where ice age iron fertilization appears to have fuelled a simple increase in export production.

## METHODS SUMMARY

The methods used in the analysis of *n*-alkanes, Fe, Ti and $^{232}$Th have been described in detail elsewhere[5]. The high-resolution records of Fe and Ti measured with the X-ray fluorescence (XRF) core-scanner were converted into quantitative concentrations using more than 400 inductively coupled plasma sector-field mass spectrometry (ICP-SFMS) measurements (Supplementary Fig. 1). Dry bulk density (DBD) was estimated for each XRF measurement interval from the $\gamma$-ray attenuation (GRA) density measurements performed in the same sediment sections as the XRF. MAR was calculated by multiplying the concentration of the different compounds by the DBD and the sedimentation rate. After error propagation, we find that the analytical component of the uncertainty is 7.8%, 8.4% and 8.6% of the final value ($1\sigma$) for Fe, Ti and *n*-alkanes MAR, respectively. The influence of age model uncertainty on the estimated MAR is evaluated in Supplementary Fig. 6. Our results show that the analytical uncertainty and the potential errors in the age model do not result in significant deviations from the estimated MAR that would affect the conclusions of this study.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 14 February; accepted 16 June 2011.**
**Published online 3 August 2011.**

1. Martin, J. H., Gordon, R. M. & Fitzwater, S. E. Iron in Antarctic waters. *Nature* **345,** 156–158 (1990).
2. Martin, J. Glacial-interglacial $CO_2$ change: the iron hypothesis. *Paleoceanography* **5,** 1–13 (1990).
3. Watson, A. J., Bakker, D. C. E., Ridgwell, A. J., Boyd, P. W. & Law, C. S. Effect of iron supply on Southern Ocean $CO_2$ uptake and implications for glacial atmospheric $CO_2$. *Nature* **407,** 730–733 (2000).
4. Kohfeld, K. E., Le Quere, C., Harrison, S. P. & Anderson, R. F. Role of marine biology in glacial-interglacial $CO_2$ cycles. *Science* **308,** 74–78 (2005).
5. Martínez-Garcia, A. *et al.* Links between iron supply, marine productivity, sea surface temperature, and $CO_2$ over the last 1.1 Ma. *Paleoceanography* **24,** PA1207, doi:10.1029/2008PA001657 (2009).
6. Sigman, D. M., Hain, M. P. & Haug, G. H. The polar ocean and glacial cycles in atmospheric $CO_2$ concentration. *Nature* **466,** 47–55 (2010).
7. Hain, M. P., Sigman, D. M. & Haug, G. H. Carbon dioxide effects of Antarctic stratification, North Atlantic Intermediate Water formation, and subantarctic nutrient drawdown during the last ice age: diagnosis and synthesis in a geochemical box model. *Glob. Biogeochem. Cycles* **24,** GB4023, doi:10.1029/2010gb003790 (2010).
8. Lambert, F. *et al.* Dust-climate couplings over the past 800,000 years from the EPICA Dome C ice core. *Nature* **452,** 616–619 (2008).
9. Wolff, E. W. *et al.* Southern Ocean sea-ice extent, productivity and iron flux over the past eight glacial cycles. *Nature* **440,** 491–496 (2006).
10. Boyd, P. W. *et al.* Mesoscale iron enrichment experiments 1993–2005: synthesis and future directions. *Science* **315,** 612–617 (2007).
11. Toggweiler, J. R. Variation of atmospheric $CO_2$ by ventilation of the ocean's deepest water. *Paleoceanography* **14,** 571–588 (1999).
12. François, R. *et al.* Contribution of Southern Ocean surface-water stratification to low atmospheric $CO_2$ concentrations during the last glacial period. *Nature* **389,** 929–935 (1997).
13. Kumar, N. *et al.* Increased biological productivity and export production in the glacial Southern Ocean. *Nature* **378,** 675–680 (1995).
14. Cassar, N. *et al.* The Southern Ocean biological response to aeolian iron deposition. *Science* **317,** 1067–1070 (2007).
15. Matsumoto, K., Sarmiento, J. L. & Brzezinski, M. A. Silicic acid leakage from the Southern Ocean: a possible explanation for glacial atmospheric $pCO_2$. *Glob. Biogeochem. Cycles* **16,** 1031, doi:10.1029/2001GB001442 (2002).
16. Brovkin, V., Ganopolski, A., Archer, D. & Rahmstorf, S. Lowering of glacial atmospheric $CO_2$ in response to changes in oceanic circulation and marine biogeochemistry. *Paleoceanography* **22,** PA4202, doi:10.1029/2006PA001380 (2007).
17. Kohler, P., Fischer, H., Munhoven, G. & Zeebe, R. E. Quantitative interpretation of atmospheric carbon records over the last glacial termination. *Global Biogeochem. Cycles* **19,** GB4020, doi:10.1029/2004GB002345 (2005).
18. Ridgwell, A. J. Implications of the glacial $CO_2$ ''iron hypothesis'' for Quaternary climate change. *Geochem. Geophys. Geosyst.* **4,** 1076, doi:10.1029/2003GC000563 (2003).
19. Jouzel, J. & Masson-Delmotte, V. Deep ice cores: the need for going back in time. *Quat. Sci. Rev.* **29,** 3683–3689 (2010).
20. Simoneit, B. R. T., Chester, R. & Eglinton, G. Biogenic lipids in particulates from the lower atmosphere over the eastern Atlantic. *Nature* **267,** 682–685 (1977).
21. Winckler, G., Anderson, R. F., Fleisher, M. Q., McGee, D. & Mahowald, N. Covariant glacial-interglacial dust fluxes in the equatorial Pacific and Antarctica. *Science* **320,** 93–96 (2008).
22. Lunt, D. J., Foster, G. L., Haywood, A. M. & Stone, E. J. Late Pliocene Greenland glaciation controlled by a decline in atmospheric $CO_2$ levels. *Nature* **454,** 1101–1105 (2008).
23. Sigman, D. M., Jaccard, S. L. & Haug, G. H. Polar ocean stratification in a cold climate. *Nature* **428,** 59–63 (2004).
24. Brierley, C. M. *et al.* Greatly expanded tropical warm pool and weakened Hadley Circulation in the early Pliocene. *Science* **323,** 1714–1718 (2009).
25. Martinez-Garcia, A., Rosell-Mele, A., McClymont, E. L., Gersonde, R. & Haug, G. H. Subpolar link to the emergence of the modern equatorial Pacific Cold Tongue. *Science* **328,** 1550–1553 (2010).
26. Honisch, B., Hemming, N. G., Archer, D., Siddall, M. & McManus, J. F. Atmospheric carbon dioxide concentration across the mid-Pleistocene transition. *Science* **324,** 1551–1554 (2009).
27. Herbert, T. D., Peterson, L. C., Lawrence, K. T. & Liu, Z. Tropical ocean temperatures over the past 3.5 million years. *Science* **328,** 1530–1534 (2010).
28. Lisiecki, L. E. & Raymo, M. E. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. *Paleoceanography* **20,** PA1003, doi:10.1029/2004PA001071 (2005).
29. Clark, P. U. *et al.* The middle Pleistocene transition: characteristics, mechanisms, and implications for long-term changes in atmospheric $pCO_2$. *Quat. Sci. Rev.* **25,** 3150–3184 (2006).
30. Cortese, G. & Gersonde, R. Plio/Pleistocene changes in the main biogenic silica carrier in the Southern Ocean, Atlantic Sector. *Mar. Geol.* **252,** 100–110 (2008).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank S. Stefer for performing the XRF scanner measurements at the University of Bremen; and I. Vöge for assistance in the ICP-SFMS analysis at the Alfred Wegener Institute for Polar and Marine Research. We thank the Integrated Ocean Drilling Program for providing the samples used in this study. This research used data acquired at the XRF Core Scanner Laboratory at the MARUM – Center for Marine Environmental Sciences, University of Bremen. Support for this work was provided by the Spanish Ministry of Science and Innovation (MICINN), the European Commission, and the Deutsche Forschungsgemeinschaft (DFG).

**Author Contributions** A.M.-G., A.R.-M. and G.H.H. designed the study. A.M.-G. performed the *n*-alkane and elemental ICP-SFMS analysis and wrote the first version of the manuscript. G.H.H. and S.L.J organized and supervised the XRF scanning at the University of Bremen. W.G. organized and supervised the ICP-SFMS elemental analysis at the Alfred Wegener Institute. All the authors contributed to the interpretation of the data and provided significant input to the final manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.M.-G (alfredo.martinez-garcia@erdw.ethz.ch).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Analysis of *n*-alkanes.** The method used for analysis of *n*-alkanes has been described in detail elsewhere[5,31,32]. Briefly, sediment samples were freeze-dried, homogenized, microwave extracted, and analysed with a Thermo Trace GC-FID. The identification of the different compounds was achieved through comparison of the chromatographic relative retention times with those of the target compounds in standards. Several selected samples were analysed by mass spectrometry to confirm peak identities and the absence of co-eluting peaks. The reproducibility of the procedure was evaluated using a homogeneous sediment standard, extracted with every batch of 14 samples. The external analytical reproducibility in the determination of *n*-alkanes in the different batches analysed was 7%.

**ICP-SFMS analysis of Fe, Ti, $^{232}$Th.** The methods used in the analysis of Fe, Ti and $^{232}$Th have been described in detail elsewhere[5]. Briefly, sediment samples were freeze-dried and digested in a pressure-assisted microwave system using an acid mixture of $HNO_3$, HCl and HF. Fe, Ti and $^{232}$Th concentrations were determined by inductively coupled plasma sector-field mass spectrometry (ICP-SFMS, Element2, Thermo Scientific). The calibration was done with dilutions of a standard solution, and Rh was used as internal standard. External reproducibility was evaluated using the NIST standard reference material 2702 ("inorganics in marine sediment"), and with each batch of samples, two procedural blanks were run. The average relative error for Fe and Ti was respectively 1.8% and 2.1% ($1\sigma$), and the external reproducibility of NIST 2702 was 5.6% and 6.5% respectively, for the batches included in Supplementary Fig. 1.

**XRF scanner measurements.** The high resolution elemental analysis of Fe and Ti was performed using an Aavatech profiling X-ray Fluorescence (XRF) core scanner at Bremen University at a 1 cm downcore resolution. The external reproducibility of the Aavatech XRF core scanner for Fe and Ti in the range of the measurements is below 2% ($1\sigma$).

**Age model.** In the interval 0–800 kyr ago, we use the age model generated by graphic correlation of the XRF Fe measurements to the ice core dust reconstruction from the EPICA project[8] (Supplementary Fig. 2). This allows us to compare the marine and continental records on similar timescales. In the interval 800–2,900 kyr ago, we use the original age model of ODP Site 1090 based on benthic $\delta^{18}$O stratigraphy[33]. In the older part of the record we have slightly modified the age model based on biostratigraphy[34,35] by aligning the *n*-alkane record to the Lisiecki and Raymo (LR04) global benthic $\delta^{18}$O stack (see Supplementary Fig. 3).

**Mass accumulation rate (MAR) calculation.** The high-resolution estimates of Fe and Ti measured with the XRF core-scanner have been converted into quantitative concentrations using more than 400 ICP-SFMS measurements. The correlation coefficients and the regression lines used for the calculation are shown in Supplementary Fig. 1. Dry bulk density (DBD) has been estimated for each XRF measurement interval from the $\gamma$-ray attenuation (GRA) density measurements[35] performed in the same sediment sections as the XRF. GRA bulk density has been converted into DBD using the regression line shown in Supplementary Fig. 1a, which is obtained from the cross-correlation of the GRA density and the available DBD discrete measurements[35]. The errors associated with the DBD are less 5% of the estimated value[36]. MARs were calculated by multiplying the concentration of the different compounds by the DBD and the sedimentation rate.

**MAR uncertainty.** In order to assess the uncertainty in the MAR of the different elements, we consider the errors of the different measurements and how they are propagated in the calculation. In order to do so, we distinguish between statistical errors, derived from the uncertainty associated with the analysis, and systematic deviations, as potentially present in the age model. Systematic deviations of the age model cannot be treated like statistical errors, as they do not meet the requirements for the underlying mathematics. After error propagation, we find that the analytical component of the uncertainty is 7.8%, 8.4% and 8.6% of the final value ($1\sigma$) for Fe, Ti and *n*-alkanes MAR respectively. The influence of age model uncertainty on the estimated MAR is evaluated in Supplementary Fig. 6. Maximum age model error envelopes of 4 kyr, 6 kyr and 15 kyr for the intervals 0–1 Myr ago, 1–3 Myr ago and 3–4 Myr ago respectively[28] do not allow for significant changes in the estimated MAR that would affect the conclusions of this study.


31. McClymont, E. L., Martinez-Garcia, A. & Rosell-Mele, A. Benefits of freeze-drying sediments for the analysis of total chlorins and alkenone concentrations in marine sediments. *Org. Geochem.* **38,** 1002–1007 (2007).
32. Kornilova, O. & Rosell-Mele, A. Application of microwave-assisted extraction to the analysis of biomarker climate proxies in marine sediments. *Org. Geochem.* **34,** 1517–1523 (2003).
33. Venz, K. A. & Hodell, D. A. New evidence for changes in Plio-Pleistocene deep water circulation from Southern Ocean ODP Leg 177 Site 1090. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **182,** 197–220 (2002).
34. Zielinski, U. & Gersonde, R. Plio-Pleistocene diatom biostratigraphy from ODP Leg 177, Atlantic sector of the Southern Ocean. *Mar. Micropaleontol.* **45,** 225–268 (2002).
35. Gersonde, R. *et al.* Southern Ocean paleoceanography. Sites 1088-1094. *Proc. ODP Init. Rep.* **177,** (1999).
36. Blum, P. *Physical Properties Handbook: A Guide to the Shipboard Measurement of Physical Properties of Deep-Sea Cores* (ODP Tech. Note 26, 1997).


©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10326

# An ancient recipe for flood-basalt genesis

Matthew G. Jackson[1] & Richard W. Carlson[2]

**Large outpourings of basaltic lava have punctuated geological time, but the mechanisms responsible for the generation of such extraordinary volumes of melt are not well known[1]. Recent geochemical evidence suggests that an early-formed reservoir may have survived in the Earth's mantle for about 4.5 billion years (ref. 2), and melts of this reservoir contributed to the flood basalt emplaced on Baffin Island about 60 million years ago[3–5]. However, the volume of this ancient mantle domain and whether it has contributed to other flood basalts is not known. Here we show that basalts from the largest volcanic event in geologic history—the Ontong Java plateau[1,6,7]—also exhibit the isotopic and trace element signatures proposed for the early-Earth reservoir[2]. Together with the Ontong Java plateau, we suggest that six of the largest volcanic events that erupted in the past 250 million years derive from the oldest terrestrial mantle reservoir. The association of these large volcanic events with an ancient primitive mantle source suggests that its unique geochemical characteristics—it is both hotter (it has greater abundances of the radioactive heat-producing elements) and more fertile than depleted mantle reservoirs—may strongly affect the generation of flood basalts.**

The discovery of a surviving portion of the early-formed, homogeneous silicate Earth that existed immediately after formation of the core—referred to as primitive mantle—would place constraints on the earliest chemical evolution of the Earth and help to clarify the means by which the Earth arrived at its present geochemical state (see, for example, refs 8–10). Most models for this primitive mantle are based on the assumption that it should have relative abundances of refractory lithophile elements similar to those of carbonaceous chondrites—the presumed building blocks of the Earth[11–13]. However, the recent discovery of small ($18 \pm 5$ parts per million, p.p.m.) differences in the $^{142}$Nd to $^{144}$Nd ratio ($^{146}$Sm decays to $^{142}$Nd with a half-life of 106 million years, Myr) between the Earth and chondrites suggests that the Earth's primitive mantle may not have chondritic relative abundances of the refractory lithophile elements[14–16].

Instead, all known modern terrestrial mantle reservoirs may have evolved from a primitive precursor with Sm/Nd ratios 4.2–7.3% higher than that of chondrites, leading to a present-day $^{143}$Nd/$^{144}$Nd of 0.51290–0.51309, which translates to a present-day $\varepsilon^{143}$Nd of +5.3 to +9.0, relative to the chondritic $^{143}$Nd/$^{144}$Nd ratio of 0.51263 (ref. 17); the stated uncertainty arises from the range of $^{142}$Nd/$^{144}$Nd found in modern terrestrial lavas and chondrites—$18 \pm 5$ p.p.m.—that propagates into uncertainty in the Sm/Nd ratio and hence into the value of the present-day $^{143}$Nd/$^{144}$Nd of the primitive precursor. If the expectation of chondritic relative abundances of refractory lithophile elements is removed, the only remaining signatures of primitive mantle are Pb-isotopic compositions on the geochron (the line in Pb-isotopic space defined by samples with constant U/Pb ratios over the Earth's age) and enrichment in the primordial isotope of helium, $^{3}$He, relative to the largely radiogenic isotope, $^{4}$He. All three (Nd, Pb and He) of the isotopic characteristics expected for a primitive terrestrial reservoir were identified in 62-Myr-old flood basalts emplaced on Baffin Island and West Greenland (BIWG)[2–5]. Employing the geochemical insights gained from BIWG, we examine some of the largest large igneous provinces (LIPs)—volcanic provinces characterized by anomalously high rates of mantle melting that represent the largest volcanic events in the Earth's history—to determine whether they are associated with a primitive (albeit non-chondritic) mantle source.

Located in the southwestern Pacific, the Ontong Java Plateau (OJP) is the largest LIP on the Earth[1,6,7]. The average $\varepsilon^{143}$Nd($t$) of these lavas[6,7] plots close to the BIWG lavas (Fig. 1) and within the range predicted for the non-chondritic primitive mantle. Excluding the most incompatible and fluid mobile elements, the OJP lavas have relatively flat primitive-mantle-normalized trace-element patterns (Fig. 2) similar to the relatively flat patterns identified in the two highest $^{3}$He/$^{4}$He lavas from BIWG (Fig. 2). The flatness of the trace-element patterns[18] led Tejada *et al.* (ref. 7) to suggest that the OJP mantle is "almost" primitive, but not actually primitive because of the higher-than-chondritic $^{143}$Nd/$^{144}$Nd in the OJP lavas.

However, the discovery of a difference in $^{142}$Nd/$^{144}$Nd between modern terrestrial rocks and chondrites suggests that the $^{143}$Nd/$^{144}$Nd measured in OJP lavas overlaps with the primitive (albeit non-chondritic) terrestrial mantle. Additionally, the 120-Myr-old early-stage (Kwaimbaita- and Kroenke-type lavas[6,7]) lavas plot near the 4.43-billion-year (Gyr)-old Pb-isotope geochron, and close to the BIWG lavas (which plot closer to the 4.5-Gyr-old geochron), an observation that is consistent with these lavas sampling an ancient mantle source (Fig. 1). Although there are minor differences in the Nd and Pb isotopic composition of the OJP and BIWG lava sources, the overlapping trace-element patterns of the two LIPs suggest an origin from compositionally similar sources whose isotopic compositions are within the range expected for an early-formed reservoir.

Owing to eruption through oceanic crust, contamination of OJP lavas by the chemically and isotopically evolved material of the continental crust does not complicate the interpretation of their mantle source to the degree seen in LIPs erupted through continental lithosphere (see Methods). In contrast, several of the largest LIPs, including the BIWG, were erupted in continental settings where assimilation of continental lithosphere can obscure the primary mantle signature of the lavas. Although continental assimilation can drive the Pb-isotopic composition of lavas towards either more or less radiogenic values, this mechanism will almost certainly lower magmatic $^{143}$Nd/$^{144}$Nd (ref. 19). To test the hypothesis that the least-contaminated LIPs that erupted in continental settings contain lavas that, like the OJP and BIWG, have Pb isotopic compositions on the geochron, we examined only the subset of LIP magmas with high $\varepsilon^{143}$Nd($t$) ($>+5.2$) that fall closest to the range suggested for a non-chondritic primitive mantle. Such lavas are least likely to have suffered from assimilation of continental crust (see Methods).

The ~180-Myr-old Karoo lavas are typical of continental LIPs in that they exhibit evidence for continental crust and lithospheric mantle contamination[20]. However, the high-MgO lavas recently discovered in the Antarctic portion of the Karoo host high $\varepsilon^{143}$Nd($t$) ratios ($\varepsilon^{143}Nd_{180\ Myr} = +7.3$ to +8.4), and these lavas plot near the Pb geochron[21] (Fig. 1).

The ~251-Myr-old Siberian Traps generally have too much of a lithospheric overprint to enable us to discern primary mantle compositions[22].

[1]Department of Earth Sciences, Boston University, 675 Commonwealth Avenue, Boston, Massachusetts 02215, USA. [2]Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington DC 20015, USA.



©2011 Macmillan Publishers Limited. All rights reserved

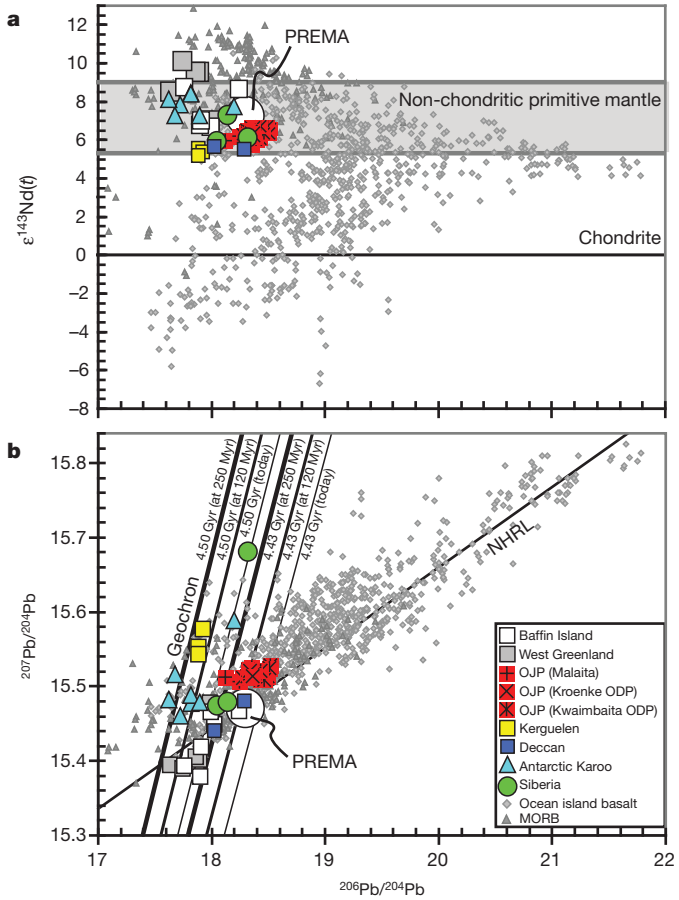



**Figure 1 | Lavas from LIPs that have Nd isotopic compositions within the range expected for a non-chondritic primitive mantle exhibit Pb isotopic compositions that plot near the geochron.** In addition to having Pb-isotopic compositions and U–Pb concentrations measured on the same samples (excluding the Deccan), all samples plotted also have paired Nd-isotopic compositions and Sm–Nd concentrations. Excluding the Pb-isotopic compositions of the Deccan lavas, all isotopic compositions are age-corrected to the time of eruption and the positions of representative age-corrected geochrons are also shown. **a**, To minimize the effects of continental assimilation, only flood basalt lavas with $\varepsilon^{143}Nd(t) \geq 5.2$ are plotted. **b**, The OJP plots closest to the 4.43-Gyr geochron at the time of eruption (120 Myr). Data for the OJP include lavas from Malaita island and ODP Leg 192 drilling sites[6,7]; late-stage Singgalo lavas are not plotted. Data sources for the other LIPs are in the text. Ocean-island-basalt and MORB data are from Georoc (http://georoc.mpch-mainz.gwdg.de/georoc/). The Northern Hemisphere reference line (NHRL) is shown[30]. $\varepsilon^{143}Nd(t) = [(^{143}Nd/^{144}Nd)_{sample}/(^{143}Nd/^{144}Nd)_{reference} - 1] \times 10{,}000$, where the isotopic ratios are calculated at the eruption age of the basalt ($t$) using the chondritic reference[17].

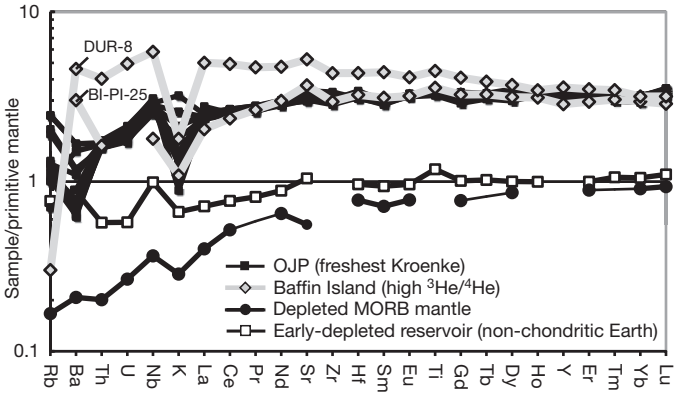



**Figure 2 | Primitive mantle[13] normalized trace-element patterns of OJP lavas compared to high $^3He/^4He$ lavas from BIWG.** For the primitive mantle composition to be consistent with past usage of this term, we use the values from ref. 13, because these have served as the model, chondrite-based, primitive mantle compositions with which the scientific community is most familiar. The lavas from Baffin Island with highest $^3He/^4He$, BI-PI-25 and DUR8, represent the normal (N) type and enriched (E) type lavas found in the BIWG LIP, respectively. Rb was not reported for BI-PI-25, and Pb concentrations were not reported for the Baffin Island or the OJP lavas[5,18]. The variability in Rb and Ba (and probably K) in OJP[18] and BIWG[2] lavas is probably due to alteration, and U has been dramatically modified by alteration in BIWG sample BI-PI-25 (and is not plotted). Only Kroenke-type OJP lavas are plotted, because they have probably experienced only olivine fractionation, whereas Kwaimbaita-type lavas are isotopically similar (Fig. 1) but are more evolved. All OJP and BIWG lavas are corrected for olivine fractionation to be in equilibrium with an olivine composition with a forsterite content of 92% (ref. 2). Only the freshest Kroenke-type lavas have been plotted, excluding the following samples (using the same filter as ref. 18): volcaniclastics and lavas with loss on ignition >0.5% and/or $K_2O/P_2O_5 > 2$. The non-chondritic primitive mantle[2] (or early depleted reservoir; see Supplementary Information) and depleted MORB mantle[15] are included.

However, high-MgO meimechites[22] have the highest $\varepsilon^{143}Nd(t)$ ratios ($\varepsilon^{143}Nd_{251\,Myr} = +5.9$ to $+7.3$) associated with the Siberian Traps. These high-MgO lavas have not been affected by the continental crust contamination that is common in most Siberian Trap lavas (Fig. 1a), and have Pb isotopic compositions near the geochron.

Basalts from the Kerguelen plateau were erupted on submerged continental crust and exhibit evidence for continental crustal assimilation. Nonetheless, ~113-Myr-old Kerguelen lavas from Ocean Drilling Program (ODP) leg 120 site 749 (ref. 23) with the highest $\varepsilon^{143}Nd(t)$ values ($\varepsilon^{143}Nd_{113\,Myr} = +5.2$ to $+5.6$) and the lowest $^{87}Sr/^{86}Sr$ values appear to be least affected by crustal assimilation, and their Pb isotopic compositions plot close to the 4.50-Gyr-old geochron (at 120 Myr).

Most formations of the ~65-Myr-old Deccan LIP are pervasively contaminated with crust and/or lithospheric mantle[24]. Lavas from the Ambenali formation have the highest $^{143}Nd/^{144}Nd$ of the Deccan basalts and are thought to best represent mantle compositions. The subset of Ambenali lavas with the highest $\varepsilon^{143}Nd(t)$ ratios ($\varepsilon^{143}Nd_{65\,Myr} = +5.5$ to $+5.6$) have Pb isotopic ratios that straddle the geochron[25].

The other test that can be applied to distinguish a primitive mantle source is high $^3He/^4He$. Unfortunately, owing to their age, post-eruptive radiogenic ingrowth of $^4He$ complicates identification of mantle source $^3He/^4He$ signatures[2,3,9].

Although many of the largest LIPs host lavas that have primitive Pb and Nd isotopic compositions, a number of large LIPs erupted in the past 250 Myr do not. This could be because of the absence of a primitive mantle component in their mantle sources or could be due to pervasive continental crust assimilation by lavas with initially primitive isotopic compositions; deconvolving the effects of crustal assimilation makes it difficult to assess the abundance of the primitive component in the mantle source of continentally erupted LIPs. Critically, however, the most frequent $^{143}Nd/^{144}Nd$ ratio in the global ocean-island basalt data set—that is, 0.5129–0.5130 (present-day $\varepsilon^{143}Nd = +5.3$ to $+9.0$; see Supplementary Information)[9]—is identical to that predicted for a non-chondritic terrestrial mantle and is also similar to the PREMA (Prevalent Mantle; see Supplementary Information) reservoir[9] (Fig. 1). This narrow range of Nd-isotopic ratios also overlaps with the $^{143}Nd/^{144}Nd$ identified in OJP and the least crustally contaminated continental LIPs (Fig. 1). If the large proportion of ocean-island-basalt lavas with present-day $\varepsilon^{143}Nd$ near $+5.3$ to $+9.0$ reflects a high proportion of non-chondritic primitive material in the mantle, then the early-formed reservoir must comprise a substantial portion of the modern terrestrial mantle. One problem with this interpretation is that many lavas with $^{143}Nd/^{144}Nd$ near 0.5130 (present-day $\varepsilon^{143}Nd = +7.2$) do not have Pb isotopic compositions that plot on the geochron (see Supplementary Information). This may be due to



©2011 Macmillan Publishers Limited. All rights reserved

incorporation of recycled oceanic crust into the primitive reservoir, which would cause the mixture to be displaced to the right of the geochron[2], owing to the high Pb concentrations and high U/Pb ratios in recycled oceanic crust[26]. The high Pb/Nd concentration ratios in subducted oceanic crust relative to other mantle reservoirs, however, would lead to only slightly lower $^{143}Nd/^{144}Nd$ in the mixture.

If the Earth's primitive mantle currently has a $^{143}Nd/^{144}Nd$ of 0.5130 (present-day $\varepsilon^{143}Nd = +7.2$), then mass-balance calculations suggest that only 45–60% of the mantle was depleted to generate the chemically complementary continental crust and depleted mid-ocean-ridge basalt (MORB) mantle[2]. Recent dynamic convection models of the Earth's mantle suggest that, over the age of the Earth, 10–25% of the Earth's convecting mantle has never melted[27]. Mantle stirring and mixing is likely to have contaminated much of the surviving primitive reservoir with recycled material. This is consistent with the hypothesis that the ocean-island-basalt source mantle is largely composed of non-chondritic primitive mantle material that has been slightly modified by the incorporation of recycled material. Therefore, the present-day mass of surviving primitive material that has escaped melting and incorporation of recycled material may be small. The fact that the Earth's major LIPs have Nd and Pb isotopic compositions close to those expected for the non-chondritic primitive reservoir suggests that these LIPs may preferentially sample the uncontaminated portions of this reservoir.

A significant fraction of the world's largest volcanic events tap a mantle source that is similar to the primitive, non-chondritic mantle reservoir discovered in high $^{3}He/^{4}He$ lavas from BIWG, and the geochemical characteristics of a non-chondritic mantle may provide insights regarding its presence in LIPs. A (non-chondritic) primitive mantle has higher concentrations of the radioactive, heat-producing elements U, Th and K (0.012, 0.046 and 159 p.p.m., respectively[2]) than does depleted MORB mantle (0.0054, 0.016 and 68.4 p.p.m.; ref. 15), though still lower than for estimates of primitive mantle based on chondritic models (0.0203, 0.0795 and 240 p.p.m.; ref. 13). This higher heat-generating capacity would make the less-depleted mantle more prone to mantle upwelling and plume generation and also cause it to melt to a greater degree. Additionally, the bulk composition of a non-chondritic mantle would be more fertile and fusible than depleted MORB mantle. Therefore, under normal melting conditions, it would undergo greater degrees of melting. These two mechanisms—a hotter and more fusible mantle source—may work together to produce the extraordinarily large volumes of melt observed at LIPs.

The current distribution of non-chondritic primitive material in the mantle is not known, but a recently discovered palaeo-geographic relationship between Phanerozoic LIPs and the large low-shear-wave-velocity provinces (LLSVPs) in the deepest mantle might provide a clue. Torsvik *et al.* (ref. 28) showed that, using a reasonable plate reconstruction, most LIPs that erupted during the past 320 Myr plot directly over one of the two hot and dense LLSVPs, one beneath Africa and one beneath the Pacific. They suggest that the two near-antipodal LLSVPs have remained stable features in the mantle for up to 540 Myr. The geographic association of LIPs at the surface with LLSVPs at depth suggest that the mantle domain represented by the LLSVPs may be the source of the LIPs. If LIPs sample a primitive non-chondritic mantle source, then the results of Torsvik *et al.* suggest that the LLSVPs are clear candidates for hosting this ancient mantle source. At about 2% of the mantle's mass[29], the two LLSVPs are sufficiently large to source all LIP volcanism over the age of the Earth (see Supplementary Information), an observation that highlights their potential as reservoirs of early primitive material in the Earth.

## METHODS SUMMARY

OJP lavas were erupted in an oceanic environment where assimilation of continental crust does not complicate the geochemical signatures of the lavas, making it possible clearly to identify their mantle source compositions. Although the OJP lavas potentially could have been contaminated by the oceanic crust through which they erupted, the relatively small difference in isotopic composition among the OJP lavas, and the low Nd and Pb concentrations of typical basaltic oceanic crust, would cause only minor isotopic shifts in the pre-contamination magmas, given reasonable amounts of contamination. Assimilation of oceanic pelagic sediments could more dramatically alter the isotopic composition of primitive OJP lavas, but such materials tend to have low $^{143}Nd/^{144}Nd$ and Pb isotope compositions plotting well to the right of the geochron. The OJP lavas display no trends consistent with the incorporation of pelagic sediments. In contrast, flood basalts erupted in continental settings are prone to contamination by assimilation of continental crust. In spite of being erupted in a continental setting, the subset of BIWG flood basalt lavas examined by ref. 2 exhibit little evidence for contamination by continental crust. In the Methods section, we examine a subset of lavas from each of four additional continental flood basalt provinces—the Antarctic portion of the Karoo flood basalt, the Siberian Traps, the Kerguelen Plateau and the Deccan—that, like the OJP and a subset of the BIWG, have suffered minimal crustal contamination and exhibit mantle source signatures predicted for a (non-chondritic) primitive mantle signature. Each of the four additional continental flood basalts have been extensively studied by others (see Methods), and the geochemical signatures associated with continental assimilation have also been thoroughly explored: the lavas that have suffered from the least continental crust assimilation also have the highest $^{143}Nd/^{144}Nd$ ($\varepsilon^{143}Nd$) ratios, and these ratios cluster near the range predicted for a non-chondritic primitive mantle.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 14 February; accepted 22 June 2011.**
**Published online 27 July 2011.**

1. Coffin, M. C. & Eldholm, O. Large igneous provinces: crustal structure, dimensions, and external consequences. *Rev. Geophys.* **32,** 1–36 (1994).
2. Jackson, M. G. *et al.* Evidence for the survival of the oldest terrestrial mantle reservoir. *Nature* **466,** 853–856 (2010).
3. Graham, D. W. *et al.* Helium isotope composition of the early Iceland mantle plume inferred from the Tertiary picrites of West Greenland. *Earth Planet. Sci. Lett.* **160,** 241–255 (1998).
4. Stuart, F. M., Lass-Evans, S., Fitton, J. G. & Ellam, R. M. High $^{3}He/^{4}He$ ratios in picritic basalts from Baffin Island and the role of a mixed reservoir in mantle plumes. *Nature* **424,** 57–59 (2003).
5. Starkey, N. A. *et al.* Helium isotopes in early Iceland plume picrites: Constraints on the composition of high $^{3}He/^{4}He$ mantle. *Earth Planet. Sci. Lett.* **277,** 91–100 (2009).
6. Tejada, M. L. G., Mahoney, J. J., Neal, C. R., Duncan, R. A. & Petterson, M. G. Basement geochemistry and geochronology of central Malaita, Solomon Islands, with implications for the origin and evolution of the Ontong Java Plateau. *J. Petrol.* **43,** 449–484 (2002).
7. Tejada, M. L. G. *et al.* in *Origin and Evolution of the Ontong Java Plateau* (eds Fitton, J. G., Mahoney, J. J., Wallace, P. J. & Saunders, A. D.) 133–150 (Geological Society of London Special Publication 229, 2004).
8. Jacobsen, S. B. & Wasserburg, G. J. The mean age of mantle and crustal reservoirs. *J. Geophys. Res.* **84,** 7411–7427 (1979).
9. Zindler, A. & Hart, S. R. Chemical geodynamics. *Annu. Rev. Earth Planet. Sci.* **14,** 493–571 (1986).
10. DePaolo, D. J. Crustal growth and mantle evolution: inferences from models of element transport and Nd and Sr isotopes. *Geochim. Cosmochim. Acta* **44,** 1185–1196 (1980).
11. Jagoutz, E. *et al.* The abundances of major, minor and trace elements in the Earth's mantle as derived from primitive ultramafic nodules. *Proc. Lunar Planet. Sci. Conf.* **X,** 203l–2050 (1979).
12. Palme, H. & O'Neill, H. S. C. in *The Mantle and Core*: *Treatise on Geochemistry* (ed. Carlson, R. W.) Vol. 2, 1–38 (Elsevier, 2003).
13. McDonough, W. F. & Sun, S. S. The composition of the Earth. *Chem. Geol.* **120,** 223–253 (1995).
14. Boyet, M. & Carlson, R. W. $^{142}Nd$ evidence for early (>4.53 Ga) global differentiation of the silicate Earth. *Science* **309,** 576–581 (2005).
15. Boyet, M. & Carlson, R. W. A new geochemical model for the Earth's mantle inferred from $^{146}Sm$–$^{142}Nd$ systematic. *Earth Planet. Sci. Lett.* **250,** 254–268 (2006).
16. Caro, G. & Bourdon, B. Non-chondritic Sm/Nd ratio in the terrestrial planets: consequences for the geochemical evolution of the mantle–crust system. *Geochim. Cosmochim. Acta* **74,** 3333–3349 (2010).
17. Bouvier, A., Vervoort, J. D. & Patchett, P. J. The Lu–Hf and Sm–Nd isotopic composition of CHUR: constraints from unequilibrated chondrites and implications for the bulk composition of terrestrial planets. *Earth Planet. Sci. Lett.* **273,** 48–57 (2008).
18. Fitton, J. G. & Goddard, M. in *Origin and Evolution of the Ontong Java Plateau* (eds Fitton, J. G., Mahoney, J. J., Wallace, P. J. & Saunders, A. D.) 151–178 (Geological Society of London Special Publication 229, 2004).
19. Carlson, R. W. Physical and chemical evidence on the cause and source characteristics of flood basalt volcanism. *Aust. J. Earth Sci.* **38,** 525–544 (1991).
20. Hawkesworth, C. J., Marsh, J. S., Duncan, A. R., Erlank, A. J. & Norry, M. J. in *Petrogenesis of the Volcanic Rocks of the Karoo Province* (ed. Erlank, A. J.) 341–354 (Geological Society of South Africa Special Publication, 1984).



©2011 Macmillan Publishers Limited. All rights reserved

21. Heinonen, J. S., Carlson, R. W. & Luttinen, A. V. Isotopic (Sr, Nd, Pb, and Os) composition of highly magnesian dikes of Vestfjella, western Dronning Maud Land, Antarctica: a key to the origins of the Jurassic Karoo large igneous province. *Chem. Geol.* **277,** 227–244 (2010).
22. Carlson, R. W., Czamanske, G., Fedorenko, V. & Ilupin, I. A comparison of Siberian meimechites and kimberlites: implications for the source of high-Mg alkalic magmas and flood basalts. *Geochem. Geophys. Geosyst.* **7,** Q11014 (2006).
23. Frey, F. A., Weis, D., Borisova, A. Y. & Xu, G. Involvement of continental crust in the formation of the Cretaceous Kerguelen Plateau: new perspectives from ODP Leg 120 Sites. *J. Petrol.* **43,** 1207–1239 (2002).
24. Peng, Z. X., Mahoney, J. J., Hooper, P. R., Harris, C. & Beane, J. E. A role for lower continental crust in flood basalt genesis? Isotopic and incompatible element study of the lower six formations of the western Deccan Traps. *Geochim. Cosmochim. Acta* **58,** 267–288 (1994).
25. Lightfoot, P. C., Hawkesworth, C. J., Devey, C. W., Rogers, N. W. & van Calsteren, P. W. C. Source and differentiation of Deccan Trap Lavas: implications of geochemical and mineral chemical variations. *J. Petrol.* **31,** 1165–1200 (1990).
26. Hofmann, A. W. & White, W. M. Mantle plumes from ancient oceanic crust. *Earth Planet. Sci. Lett.* **57,** 421–436 (1982).
27. Brandenburg, J. P., Hauri, E. H., van Keken, P. E. & Ballentine, C. J. A multiple-system study of the geochemical evolution of the mantle with force-balanced plates and thermochemical effects. *Earth Planet. Sci. Lett.* **276,** 1–13 (2008).
28. Torsvik, T. H., Burke, K., Steinberger, B., Webb, S. J. & Ashwal, L. D. Diamonds sampled by plumes from the core–mantle boundary. *Nature* **466,** 352–355 (2010).
29. Burke, K. Plate tectonics, the Wilson Cycle, and mantle plumes: geodynamics from the top. *Annu. Rev. Earth Planet. Sci.* **39,** 1–29 (2011).
30. Hart, S. R. A large-scale isotope anomaly in the southern hemisphere mantle. *Nature* **309,** 753–757 (1984).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank J. Day, S. Hart, N. Shimizu, S. Shirey, J. Mahoney and M. Kurz for discussions and B. Hanan and G. Fitton for their detailed review comments. M.J. acknowledges Boston University start-up funds and the Ocean Sciences Section of the National Science Foundation that supported this work.

**Author Contributions** M.J. and R.C. contributed equally to the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.G.J. (jacksonm@bu.edu).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

To establish a baseline for the composition of melts sampling a (non-chondritic) primitive mantle reservoir, this Letter relies on the conclusions drawn from the OJP data set, in large part to avoid the issue of the sensitivity of continental flood basalts to contamination by continental crust during emplacement and eruption. The OJP was erupted in an oceanic environment, and there is no evidence for continental fragments in the OJP. Although we cannot rule out small degrees of contamination from the oceanic lithosphere, the Pb-isotopic composition of OJP lavas is not at all consistent with a strong oceanic crust or sediment overprint. The tight clustering of the OJP Pb-isotopic data—which were measured on samples collected from Malaita and multiple drill core locations scattered over hundreds of kilometres across the plateau—is remarkable and argues against any systematic offset of the data caused by oceanic crust or sediment contamination. An important point regarding the OJP lavas is that we did not filter the data on the basis of any isotopic parameter. Figure 1 shows all available OJP lavas (excluding late-stage Singgalo lavas) that have both paired U–Pb concentrations and Pb-isotopic measurements and paired Sm–Nd concentrations and Nd-isotopic measurements.

In contrast, a specific isotopic filter was applied to the continental flood basalts, because the isotopic compositions of such lavas are often overprinted by continental crust. In selecting samples with the most radiogenic Nd-isotopic compositions, we are selecting samples with the lowest crustal input and it is those lavas that have Pb-isotopic compositions that plot close to the geochron. Although we cannot rule out the possibility that these basalts obtained their Pb-isotopic compositions through a haphazard mixture of continental crust and variable mantle sources, it is worth noting that all of the samples shown in Fig. 1—that were selected based solely on Nd-isotopic compositional characteristics—also plot exactly where expected in Pb-isotopic space for a melt of ancient primitive mantle. Below, we present additional geochemical evidence from the literature that supports our contention that continental flood basalts—including BIWG, the Kerguelen Plateau, Siberian Traps, Deccan Traps and the Antarctic portion of the Karoo—with the highest $^{143}Nd/^{144}Nd$ have suffered from the least crustal contamination and best represent mantle source compositions.

**BIWG flood basalt lavas.** Continental crust contamination had already been ruled out for a suite of lavas from BIWG examined by ref. 2. Larsen and Pederson (ref. 31) used trace-element indicators sensitive to continental crust assimilation to show that West Greenland lavas with the least crustal contamination have radiogenic $\varepsilon^{143}Nd(t)$ values and plot in a specific region of Pb-isotopic space. An evaluation of this data by ref. 2 showed that the region of Pb-isotopic space occupied by the least contaminated Baffin Island lavas corresponds with the 4.55–4.45-Gyr terrestrial geochrons.

**Kerguelen Plateau basalts.** Although originally viewed as an oceanic plateau, continental fragments have been identified in the Kerguelen Plateau using seismic methods (ref. 32). Clasts of garnet-biotite gneiss recovered from Elan Bank contain zircons with ages ranging from 534 to 2,547 Myr (ref. 33), clearly indicative of an old crustal basement beneath portions of the Kerguelen Plateau. Additionally, studies in the area have found geochemical evidence for continental assimilation in Kerguelen basalts, and workers in this area have spent considerable efforts showing which basalts are and are not contaminated in this province[23,34–37]. For example, Kerguelen lavas that have been contaminated with continental crust tend to have low Nb/La (for example, ODP leg 120 site 747, ODP leg 183 site 1137) or high $^{87}Sr/^{86}Sr$ (>0.7045; for example, ODP leg 119 site 738, ODP leg 120 sites 747 and 750, and ODP leg 183 site 1137). In contrast, a subset of the lavas from ODP leg 120 site 749 have low $^{87}Sr/^{86}Sr$ (<0.704) and Nb/La ratios (0.9–1.0) that are similar to oceanic lavas and low $^{87}Sr/^{86}Sr$ ratios[23]. Site 749 lavas are divided into two distinct groups: group 1 has lower Nb/La (0.80–0.85) and higher $^{87}Sr/^{86}Sr$ (>0.704) than group 2, which has higher Nb/La (0.9–1.0) and lower $^{87}Sr/^{86}Sr$. The site 749 lavas from group 1, with the lower Nb/La and higher $^{87}Sr/^{86}Sr$, are interpreted to have suffered from more assimilation of continental crust than the group 2 lavas. We note that the group 2 lavas from site 749 basalts, which we consider to be the least contaminated by continental crust, have the highest $\varepsilon^{143}Nd(t)$ in the Kerguelen Plateau suite, an observation that supports our contention that continental LIPs with the most radiogenic Nd-isotopic compositions best represent the mantle source in geochemically complex continental flood basalt provinces. These lavas also have Pb-isotopic compositions that plot close to the geochron.

**Deccan flood basalt lavas.** In the Deccan, the Ambenali formation flows are thought to best represent mantle compositions[24,38,39]. Several studies have documented the nature of crustal contamination in the Deccan and show that the least-contaminated members of the Ambenali are closest to mantle source compositions[24,38,39]. For example, ref. 38 used Ba/Y and Ba/Zr (where both ratios are high in continental crust) to identify lavas that have suffered from crustal contamination, and they found that lavas with the lowest Ba/Y and Ba/Zr are found in the Ambenali formation. Additionally, ref. 39 used Ba/Ti to evaluate continental assimilation (where high Ba/Ti is high in continental crust), and they found Ambenali lavas to have low Ba/Ti ratios. Ref. 39 also considered radiogenic Sr isotopic ratios to be a signature of crustally contaminated Deccan lavas, and although Ambenali lavas tend to have the lowest $^{87}Sr/^{86}Sr$ in the Deccan, some of the Ambenali lavas have elevated $^{87}Sr/^{86}Sr$—with correspondingly low $\varepsilon^{143}Nd(t)$—that are considered to be crustally contaminated. Like ref. 39, we exclude such lavas as having a mantle source composition. In summary, the least-contaminated units of the Ambenali formation, which is considered to be the least contaminated in the Deccan, also have the highest $\varepsilon^{143}Nd(t)$, and these lavas have Pb-isotopic compositions that plot close to the geochron.

**Siberian flood basalt.** The ~251-Myr-old Siberian Traps are generally too contaminated by continental crust assimilation to allow us to evaluate primary mantle compositions[22,40]. However, a recent study showed that the Siberian meimichites have the highest $\varepsilon^{143}Nd(t)$ of any Siberian flood basalt and have trace-element characteristics inconsistent with crustal contamination[22]. Together with having the highest $\varepsilon^{143}Nd(t)$, the meimechites also have the lowest Th/Ta in the flood basalt province. This is in contrast to the crustally contaminated Siberian flood basalt lavas that have high Th/Ta (continental crust has high Th/Ta) coupled with low $\varepsilon^{143}Nd(t)$ (ref. 22). The Siberian meimechites also have Os-isotopic compositions inconsistent with any significant crustal input[22]. Like the Deccan and Kerguelen flood basalts, the least crustally contaminated Siberian flood basalt lavas have the highest $\varepsilon^{143}Nd(t)$, and these lavas have Pb-isotopic compositions that plot close to the geochron.

**Antarctic Karoo flood basalts.** Using a data set similar to the Siberian lavas from ref. 22, similar arguments were used to argue for a lack of contamination in the picritic lavas in the Antarctic portion of the Karoo flood basalt province[21]. The lavas that are considered to be least contaminated by continental crust exhibit Os-isotopic signatures that are thought to reflect mantle signatures (age-corrected $^{188}Os/^{188}Os = 0.1257–0.1286$), while the crustally contaminated lavas have higher $^{187}Os/^{188}Os$ (0.1297–0.1426), consistent with continental crust assimilation. The crustally contaminated lavas also have higher $^{87}Sr/^{86}Sr$ (≥0.7036 after age correction) than the least crustally contaminated lavas (<0.7036). The least-contaminated Antarctic Karoo lavas also have the highest $\varepsilon^{143}Nd(t)$, and their Pb-isotopic compositions plot close to the geochron.


31. Larsen, L. M. & Pedersen, A. K. Petrology of the Paleocene picrites and flood basalts on Disko and Nuussuaq, West Greenland. *J. Petrol.* **50,** 1667–1711 (2009).
32. Operto, S. & Charvis, P. Deep structure of the southern Kerguelen Plateau (southern Indian Ocean) from ocean bottom seismometer wide-angle seismic data. *J. Geophys. Res.* **101,** 25077–25103 (1996).
33. Nicolaysen, K. *et al.* Provenance of Proterozoic garnet-biotite gneiss recovered from Elan Bank, Kerguelen Plateau, southern Indian Ocean. *Geology* **29,** 235–238 (2001).
34. Ingle, S., Weis, D., Doucet, S. & Mattielli, N. Hf isotope constraints on mantle sources and shallow-level contaminants during Kerguelen hotspot activity since ~120 Ma. *Geochem. Geophys. Geosyst.* **4,** (2003).
35. Mahoney, J. *et al.* Geochemical characteristics of lavas from Broken Ridge, the Naturaliste Plateau and Southernmost Kerguelen Plateau: early volcanism of the Kerguelen hotspots. *Chem. Geol.* **120,** 315–345 (1995).
36. Ingle, S., Weis, D., Scoates, J. S. & Frey, F. A. Relationship between the early Kerguelen plume and continental flood basalts of the paleo-Eastern Gondwanan margins. *Earth Planet. Sci. Lett.* **197,** 35–50 (2002).
37. Storey, M. *et al.* Lower Cretaceous volcanic rocks on continental margins and their relationship to the Kerguelen Plateau. *Proc. ODP Sci. Res.* **120,** 33–53 (1992).
38. Lightfoot, P. & Hawkesworth, C. Origin of Deccan Trap lavas: evidence from combined trace element and Sr-, Nd- and Pb-isotope studies. *Earth Planet. Sci. Lett.* **91,** 89–104 (1988).
39. Peng, Z. X. & Mahoney, J. J. Drillhole lavas from the northwestern Deccan Traps, and the evolution of R union hotspot mantle. *Earth Planet. Sci. Lett.* **134,** 169–185 (1995).
40. Fedorenko, V. A. *et al.* Petrogenesis of the flood-basalt sequence at Noril'sk, North Central Siberia. *Int. Geol. Rev.* **38,** 99–135 (1996).


©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10249

# Using the *Acropora digitifera* genome to understand coral responses to environmental change

Chuya Shinzato[1]*, Eiichi Shoguchi[1]*, Takeshi Kawashima[1]*, Mayuko Hamada[1]*, Kanako Hisata[1], Makiko Tanaka[1], Manabu Fujie[2], Mayuki Fujiwara[1], Ryo Koyanagi[1], Tetsuro Ikuta[1], Asao Fujiyama[3], David J. Miller[4] & Nori Satoh[1]

**Despite the enormous ecological and economic importance of coral reefs, the keystone organisms in their establishment, the scleractinian corals, increasingly face a range of anthropogenic challenges including ocean acidification and seawater temperature rise[1–4]. To understand better the molecular mechanisms underlying coral biology, here we decoded the approximately 420-megabase genome of *Acropora digitifera* using next-generation sequencing technology. This genome contains approximately 23,700 gene models. Molecular phylogenetics indicate that the coral and the sea anemone *Nematostella vectensis* diverged approximately 500 million years ago, considerably earlier than the time over which modern corals are represented in the fossil record (~240 million years ago)[5]. Despite the long evolutionary history of the endosymbiosis, no evidence was found for horizontal transfer of genes from symbiont to host. However, unlike several other corals, *Acropora* seems to lack an enzyme essential for cysteine biosynthesis, implying dependency of this coral on its symbionts for this amino acid. Corals inhabit environments where they are frequently exposed to high levels of solar radiation, and analysis of the *Acropora* genome data indicates that the coral host can independently carry out *de novo* synthesis of mycosporine-like amino acids, which are potent ultraviolet-protective compounds. In addition, the coral innate immunity repertoire is notably more complex than that of the sea anemone, indicating that some of these genes may have roles in symbiosis or coloniality. A number of genes with putative roles in calcification were identified, and several of these are restricted to corals. The coral genome provides a platform for understanding the molecular basis of symbiosis and responses to environmental changes.**

Coral reefs are estimated to harbour around one third of all described marine species[6], and their productivity supports around one quarter of marine fisheries, but declines in coral abundance and wholesale loss of reef habitats are one of the most pressing environmental issues of our time. The major architects of coral reefs, the scleractinian corals, are anthozoan cnidarians that form obligate endosymbioses with photosynthetic dinoflagellates of the genus *Symbiodinium* (Fig. 1b). The symbionts confer on the coral holobiont the ability to fix $CO_2$ and to deposit the massive aragonite (a form of calcium carbonate) skeletons that distinguish reef-building corals from other anthozoans such as sea anemones. The association is fragile, however, and collapses under stress. Despite the ecological and economic significance of corals, the molecular mechanisms underlying much of coral biology—including stress responses and disease—remain unclear, but it is clear that corals retain much of the complex gene repertoire of the ancestral metazoan[7]. To address the lack of molecular data for reef-building corals, we determined the whole-genome sequence of *A. digitifera* (Fig. 1a–h), a dominant species on Okinawan reefs. Not only are *Acropora* species the dominant reef-building corals of the Indo-Pacific, but they are also among the most sensitive of corals to increased seawater temperatures[8].

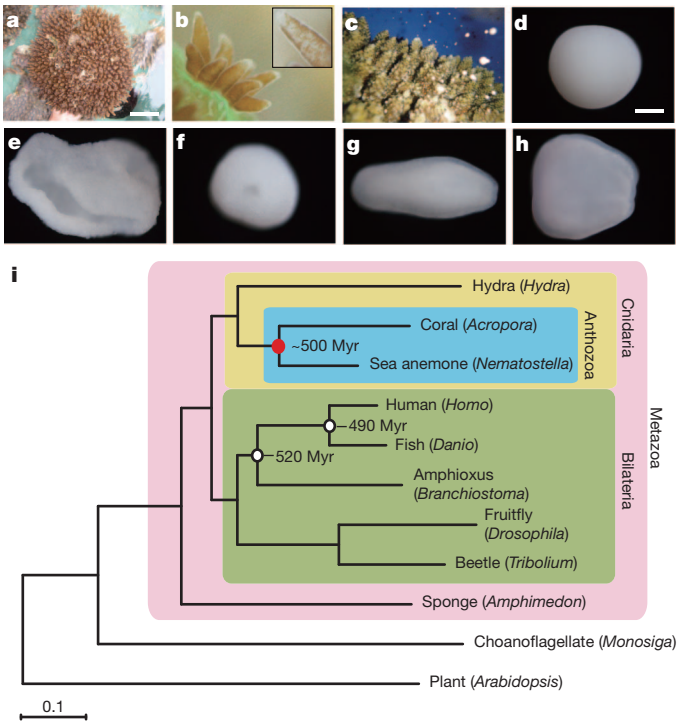



**Figure 1 | The coral *Acropora digitifera* and an early occurrence of corals on Earth. a**, The colony the genome of which was sequenced in the present study. This colony is maintained in aquarium culture at the Sesoko Station, University of the Ryukyus, and is thus available for further investigation of the genome. Scale bar, 10 cm. **b**, Polyps of the coral showing the presence of symbiotic dinoflagellates (*Symbiodinium* sp.) (inset, enlargement). **c**, Natural spawning of the coral. **d**–**h**, Eggs, embryos, larva and primary polyp of *A. digitifera*, from which messenger RNA was extracted for transcriptome analyses. Scale bar, 200 μm. **d**, Fertilized egg; **e**, blastula at the prawn chip stage; **f**, gastrula; **g**, planula larva; and **h**, primary polyp. **i**, Molecular phylogeny of corals. 94,200 aligned amino acid positions of proteins encoded by 422 genes were obtained from the sponge *Amphimedon queenslandica*, from the cnidarians *A. digitifera*, *Nematostella vectensis* and *Hydra magnipapillata*, and from the triploblasts *Tribolium castaneum*, *Drosophila melanogaster*, *Branchiostoma floridae*, *Danio rerio* and *Homo sapiens*. The sequences were analysed using maximum likelihood methods, with the plant *Arabidopsis thaliacia* and the choanoflagellate *Monosiga brevicollis* serving as outgroups. The scale bar represents 0.1 expected substitutions per site in the aligned regions. The topology was supported by 100% bootstrap value. Approximate divergent times of the occurrence of basal chordates and divergence of vertebrates lineages are shown. This analysis indicates a deeper divergence of *Acropora* and *Nematostella*, approximately 500 million years (Myr) ago.

[1]Marine Genomics Unit, Okinawa Institute of Science and Technology Promotion Corporation, Onna, Okinawa 904-0412, Japan. [2]DNA Sequencing Center Section, Okinawa Institute of Science and Technology Promotion Corporation, Onna, Okinawa 904-0412, Japan. [3]National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan. [4]ARC Centre of Excellence for Coral Reef Studies and School of Pharmacy and Molecular Sciences, James Cook University, Townsville, Queensland 4811, Australia.
*These authors contributed equally to this work.



©2011 Macmillan Publishers Limited. All rights reserved

On the basis of flow cytometry, the *A. digitifera* genome is approximately 420 Mbp (Supplementary Figs 1 and 2) and is therefore similar in size to that of the sea anemone *Nematostella*[9]. Sperm from a single colony served as the source of DNA for sequencing using a combination of Roche 454 GS-FLX[10] and Illumina Genome Analyser IIx (GAIIx)[11] methods. The genome was sequenced to approximately 151-fold coverage (Supplementary Table 1), enabling the generation of an assembly comprising a total of 419 Mbp (Supplementary Tables 2–5; contig N50 = 10.7 kbp and scaffold N50 = 191.5 kbp; Supplementary Fig. 3). The genome is approximately 39% G+C (Supplementary Fig. 4), and contains 23,668 predicted protein-coding loci (Supplementary Table 6). Transposable elements occupy approximately 12.9% of the genome (Supplementary Table 7). The coral gene set is comparable in size and composition with those of *Nematostella vectensis*[9] and *Hydra magnipapillata*[12] (Supplementary Tables 6, 8 and 9). The genome browser is accessible at http://marinegenomics.oist.jp/acropora_digitifera (Supplementary Fig. 5). Approximately 93% of the *A. digitifera* genes have matches in other metazoans (Supplementary Fig. 6a), and of these, 11% have clear homology only among expressed sequence tag (EST) data from corals[13] (Supplementary Fig. 6b), suggesting the presence of a considerable number of coral-specific genes.

Corals are morphologically very similar to sea anemones, but their evolutionary origins are obscure. Reef-building Scleractinia first appeared in the fossil record in the mid-Triassic (approximately 240 million years ago)[5], but were already highly diversified, suggesting much earlier origins. The availability of fully sequenced genomes for three cnidarians—*Acropora* (the present study), *Nematostella*[9] and *Hydra*[12]— allowed the estimation of the depth of the divergence between corals and other metazoans. Molecular phylogenetic analyses based on an alignment of 94,200 amino acid positions suggest a divergence time of *Acropora* and *Nematostella* between 520 to ~490 million years ago (the late Cambrian or early Ordovician) (Fig. 1i). The implied earlier origin of Scleractinia indicates that corals have persisted through previous periods of major environmental change, including the mass extinction event at the Permian/Triassic boundary, when global $CO_2$ and temperature were much higher than at present. However, whereas the Scleractinia as a lineage has persisted on evolutionary time scales, whether modern coral reefs can adapt to rapid environmental change on ecological time scales is a very different question.

The obligate endosymbiosis of corals dates at least from the mid-Triassic, and the longevity of this association might therefore be expected to have resulted in changes within the coral genome. We were unable to find any *Symbiodinium* DNA sequences in the coral genome, hence there is as yet no evidence for horizontal gene transfer from symbiont to host (Supplementary Fig. 6). However, comparative analyses indicated that, in the case of *Acropora*, the coral host might be metabolically dependent on the symbiont. Using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database[14], the metabolic repertoire of *Acropora* was compared to that of its non-symbiotic relative, the sea anemone *Nematostella* (Supplementary Table 10), leading to the identification of an apparent metabolic deficiency in *Acropora*. The biosynthesis of cysteine from homocysteine and/or serine requires the activities of two enzymes, cystathionine β-synthase (Cbs) and cystathionase (cystathionine γ-lyase; Cth) (Table 1). Whereas we were able to identify genes encoding the latter in both *A. digitifera* and *Nematostella*, the former could not be identified in *Acropora* despite a clear match being present in *Nematostella* (Supplementary Fig. 7). Although extensive transcriptomic data are available for various *Acropora* spp[13], we could find no evidence for a *Cbs* transcript in any of these. Moreover, whereas a polymerase chain reaction (PCR) strategy confirmed the presence of *Cbs* in some other corals, *Galaxea fascicularis*, *Favites chinenis*, *Favia lizardensis* and *Ctenactis echinata*, no amplification products could be obtained for two different *Acropora* species (Table 1 and Supplementary Fig. 8). Although the analyses presented here do not rigorously exclude the presence of Cbs activity in *Acropora*,

**Table 1 | The presence or absence of a gene encoding Cbs for L-cysteine biosynthesis**

| | L-Homocysteine + L-Serine | *Cbs* → | L-Cystathionine | *Cth* → | L-Cysteine |
|---|---|---|---|---|---|
| *Hydra magnipapillata* | | Yes* | | Yes | |
| *Nematostella vectensis* | | Yes* | | Yes | |
| Complexa | | | | | |
| *Acropora digitifera* | | No† | | Yes | |
| *Acropora tenuis* | | No‡ | | ND | |
| *Acropora millepora* | | No§ | | Yes | |
| *Galaxea fascicularis* | | Yes‡ | | ND | |
| Robusta | | | | | |
| *Montastraea faveolata* | | Yes§ | | Yes | |
| *Favia lizardensis* | | Yes‡ | | ND | |
| *Favites chinensis* | | Yes‡ | | ND | |
| *Ctenactis echinata* | | Yes‡ | | ND | |

ND, not determined.
* Supported by sequenced genome and EST analyses.
† Supported by sequenced genome, EST and PCR amplification of genomic DNA.
‡ Supported by PCR amplification of genomic DNA.
§ Supported by EST analyses.

they raise the intriguing possibility of a metabolic basis for the obligate nature of symbiosis in *Acropora*; differences in dependency could potentially explain not only the phenomenon of symbiont selectivity, but also the high sensitivity of *Acropora* to environmental challenges.

Reef-building corals typically inhabit shallow and relatively clear tropical waters and are therefore constantly exposed to high levels of ultraviolet irradiation. As corals are particularly susceptible to bleaching when exposed to both raised temperatures and high solar radiation[2,4], one intriguing question is how corals protect themselves against ultraviolet damage. Photo-protective compounds, such as the mycosporine amino acids (MAAs), have been isolated from corals[15,16] but, because similar compounds have been identified in algae, the sources of these compounds were unknown. Recently a short (four-step) pathway encoded by a gene cluster (DHQS-like, *O-MT*, ATP-grasp and NRPS-like) (Fig. 2 and Supplementary Figs 9–12) has been demonstrated to be both necessary and sufficient in the cyanobacterium *Anabaena variabilis* to convert pentose-phosphate metabolites to shinorine, a photo-protective MAA[17]. Scanning the available whole-genome data allowed us to identify clear homologues of all four members of the cyanobacterial shinorine gene cluster in both *A. digitifera* and *N. vectensis* (Fig. 2), indicating that both the coral and the sea anemone have the ability to carry out *de novo* synthesis of ultraviolet-protective compounds. Hence, MAA synthesis in corals and other cnidarians is not symbiont dependent.

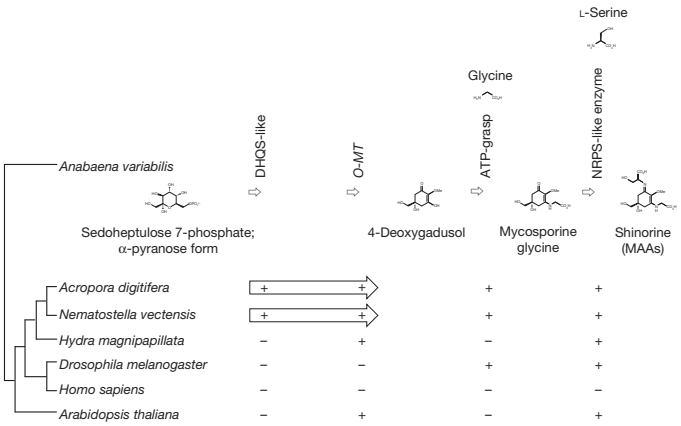



**Figure 2 | The genes required for the biosynthesis of shinorine are present in anthozoan cnidarians.** Top, the organization of the gene cluster involved in the biosynthetic pathway of the photo-protective molecule shinorine, a mycosporine-like amino acid, in the cyanobacterium *Anabaena variabilis*. Bottom, the presence of corresponding genes in various organisms is indicated (+). The *Acropora* and *Nematostella* genomes contain homologues of each of the four genes, in which DHQS-like and *O-MT* are fused with each other.



©2011 Macmillan Publishers Limited. All rights reserved

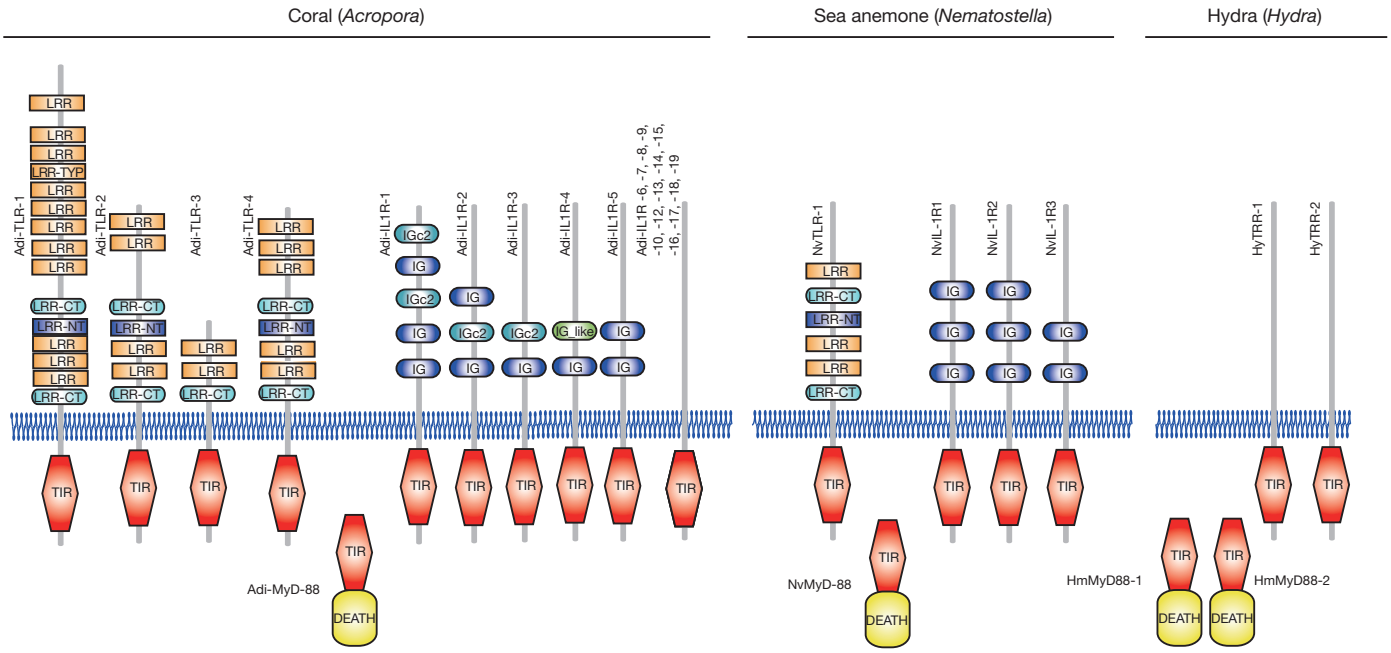



**Figure 3 | Repertoires of TIR-domain-containing proteins of three cnidarians.** The schematic representation of the domain structures of TIR-domain-containing proteins identified in *A. digitifera*, alongside the corresponding complements from *Nematostella vectensis* and *Hydra magnipapillata*. The repertoire of Toll/TLR, IL-1R-like and TIR-only proteins is significantly more complex in the case of *A. digitifera* than in *N. vectensis* or *H. magnipapillata*. TIR, TIR domain. DEATH, DEATH domain. IG and IGc2, Ig domain. LRR, LRY-TRY, LRR-CT and LRR-NT, leucine-rich repeats.

Surveys of *Acropora* for genes associated with innate immunity[18], apoptosis[19] and autophagy[19] indicate not only the complexity of these systems in *Acropora* (Supplementary Figs 13–23), but also that the coral innate immune repertoire is more sophisticated than that of *Nematostella*. For example, whereas a single canonical Toll/TLR protein is present in *N. vectensis*[18], the *Acropora* genome encodes at least four such molecules, as well as five IL-1R-related proteins and a number of TIR-only proteins (Fig. 3). Likewise, the *Acropora* repertoire of NACHT/NB-ARC domains, which are characteristic of primary intracellular pattern receptors[20], is again highly complex: an order of magnitude more NACHT/NB-ARC domains are present in coral than in other animals (Supplementary Table 11), and some of these cnidarian proteins have novel domain structures (Supplementary Fig. 23b). In terms of the apparent expansion and divergence of NACHT-encoding genes, the coral resembles amphioxus[21], the sea urchin[22] and angiosperms[23]. The greater complexity of the coral innate immunity network may in part reflect adaptations associated with the symbiotic state and coloniality.

The coral repertoire of genes with predicted roles in skeleton deposition is of particular interest given the likely impact of ocean acidification resulting from rising atmospheric $CO_2$ on coral calcification. Surveys of the *Acropora* genome for specific groups of proteins associated with calcification, including the eukaryotic-type carbonic anhydrases[24] are given in Supplementary Table 12. In general, the soluble fraction of the organic matrix in scleractinian corals is very rich in acidic amino acids, and has a particularly high aspartic acid composition[25]. A number of candidate organic-matrix proteins were identified in *Acropora* (Supplementary Fig. 24). For several of these, orthologues could be identified in *A. millepora* and/or *A. palmata* but only one of these (Adi-SAP6) was found in other coral species (Supplementary Table 13). Galaxins, first purified from the coral *Galaxea fascicularis*, are unique to corals and are the only coral skeletal matrix protein for which the complete primary structure has been determined[26]. However, galaxin possesses neither acidic regions (the fraction of Asp+Asn in the galaxin is 9.7%) nor obvious $Ca^{2+}$-binding domains[26]. Four genes encoding galaxin-related proteins were identified in the *A. digitifera* genome (Supplementary Fig. 25), including two likely *A. digitifera* homologues of Gfa-galaxin.

Here we decoded the 420-Mbp genome of the reef-building coral *Acropora digitifera*, with the aim of providing a platform for understanding the molecular basis of symbiosis and responses to environmental change. Some of the main findings are: (1) a relatively deep divergence of the lineage leading to the reef-building corals; (2) although we could find no evidence for horizontal gene transfer from symbiont to coral despite the long evolutionary history of the association, *Acropora* may have lost a gene essential for cysteine biosynthesis and thus be metabolically dependent on its symbionts; (3) the coral host has the ability to independently carry out *de novo* synthesis of the MAA family of photoprotective compounds; (4) the innate immune repertoire of coral is highly complex in comparison with the non-symbiotic and solitary sea anemone *Nematostella*; and (5) a number of coral-specific gene families are likely to have evolved in the context of calcification. These data also provide a basis for systems biology approaches to understanding the establishment, function and collapse of coral symbioses. If and when a whole-genome sequence becomes available for the dinoflagellate symbiont of corals *Symbiodinium* sp. (zooxanthellae), these resources will together provide additional perspectives on the symbiosis and a powerful resource for understanding the response of the holobiont to environmental stresses such as raised seawater temperatures or ocean acidification.

## METHODS SUMMARY

Sperm DNA obtained from a single colony of the coral *Acropora digitifera* was used for genome sequencing by Roche 454 GS-FLX[10] and Illumina Genome Analyser IIx (GAIIx)[11]. The 454 shotgun and paired-end reads were assembled *de novo* by GS *De novo* Assembler version 2.3 (Newbler, Roche)[10], and subsequent scaffolding was performed by SOPRA[27] and SSPACE[28] using the Illumina mate-pair information. Transcriptome analysis was also performed. A set of gene model predictions (the *A. digitifera* Gene Model v. 1) was generated mainly by AUGUSTUS[29], and a genome browser has been established using the Generic Genome Browser (GBrowser) 2.17. The annotation and identification of *Acropora* genes were performed by three approaches, individual methods or combinations of the methods:



©2011 Macmillan Publishers Limited. All rights reserved

reciprocal BLAST analyses, screening the gene models against the Pfam database[30] and phylogenetic analyses.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 7 January; accepted 3 June 2011.**
**Published online 24 July 2011.**

1. Hughes, T. P. *et al.* Climate change, human impacts, and the resilience of coral reefs. *Science* **301,** 929–933 (2003).
2. Hoegh-Guldberg, O. *et al.* Coral reefs under rapid climate change and ocean acidification. *Science* **318,** 1737–1742 (2007).
3. Carpenter, K. E. *et al.* One-third of reef-building corals face elevated extinction risk from climate change and local impacts. *Science* **321,** 560–563 (2008).
4. Weis, V. M. Cellular mechanisms of cnidarian bleaching: stress causes the collapse of symbiosis. *J. Exp. Biol.* **211,** 3059–3066 (2008).
5. Stanley, G. D. Jr & Fautin, D. G. Paleontology and evolution. The origins of modern corals. *Science* **291,** 1913–1914 (2001).
6. Wilkinson, C. *Status of Coral Reefs of the World* (Australian Institute of Marine Studies, 2004).
7. Miller, D. J., Ball, E. E. & Technau, U. Cnidarians and ancestral genetic complexity in the animal kingdom. *Trends Genet.* **21,** 536–539 (2005).
8. Loya, Y. *et al.* Coral bleaching: the winners and the losers. *Ecol. Lett.* **4,** 122–131 (2001).
9. Putnam, N. H. *et al.* Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* **317,** 86–94 (2007).
10. Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437,** 376–380 (2005).
11. Bentley, D. R. Whole-genome re-sequencing. *Curr. Opin. Genet. Dev.* **16,** 545–552 (2006).
12. Chapman, J. A. *et al.* The dynamic genome of *Hydra*. *Nature* **464,** 592–596 (2010).
13. Hemmrich, G. & Bosch, T. C. Compagen, a comparative genomics platform for early branching metazoan animals, reveals early origins of genes regulating stem cell differentiation. *Bioessays* **30,** 1010–1018 (2008).
14. Kanehisa, M. *et al.* KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **36,** D480–D484 (2008).
15. Shick, J. M. & Dunlap, W. C. Mycosporine-like amino acids and related Gadusols: biosynthesis, acumulation, and UV-protective functions in aquatic organisms. *Annu. Rev. Physiol.* **64,** 223–262 (2002).
16. Rastogi, R. P. *et al.* Photoprotective compounds from marine organisms. *J. Ind. Microbiol. Biotechnol.* **37,** 537–558 (2010).
17. Balskus, E. P. & Walsh, C. T. The genetic and molecular basis for sunscreen biosynthesis in cyanobacteria. *Science* **329,** 1653–1656 (2010).
18. Miller, D. J. *et al.* The innate immune repertoire in cnidaria—ancestral complexity and stochastic gene loss. *Genome Biol.* **8,** R59 (2007).
19. Dunn, S. R., Schnitzler, C. E. & Weis, V. M. Apoptosis and autophagy as mechanisms of dinoflagellate symbiont release during cnidarian bleaching: every which way you lose. *Proc. R. Soc. Lond. B* **274,** 3079–3085 (2007).
20. Meylan, E., Tschopp, J. & Karin, M. Intracellular pattern recognition receptors in the host response. *Nature* **442,** 39–44 (2006).
21. Huang, S. *et al.* Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. *Genome Res.* **18,** 1112–1126 (2008).
22. Hibino, T. *et al.* The immune gene repertoire encoded in the purple sea urchin genome. *Dev. Biol.* **300,** 349–365 (2006).
23. Leister, D. Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance gene. *Trends Genet.* **20,** 116–122 (2004).
24. Jackson, D. J., Macis, L., Reitner, J., Degnan, B. M. & Wörheide, G. Sponge paleogenomics reveals an ancient role for carbonic anhydrase in skeletogenesis. *Science* **316,** 1893–1895 (2007).
25. Sarashina, I. & Endo, K. Skeletal matrix proteins of invertebrate animals: comparative analysis of their amino acid sequences. *Paleontological Res.* **10,** 311–336 (2006).
26. Fukuda, I. *et al.* Molecular cloning of a cDNA encoding a soluble protein in the coral exoskeleton. *Biochem. Biophys. Res. Commun.* **304,** 11–17 (2003).
27. Dayarian, A., Michael, T. P. & Sengupta, A. M. SOPRA: scaffolding algorithm for paired reads via statistical optimization. *BMC Bioinformatics* **11,** 345 (2010).
28. Boetzer, M. *et al.* Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27,** 578–579 (2011).
29. Stanke, M., Diekhans, M., Baertsch, R. & Haussler, D. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. *Bioinformatics* **24,** 637–644 (2008).
30. Finn, R. D. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **38,** D211–D222 (2010).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We acknowledge A. Iguchi, K. Sakai and all other staff members of the Sesoko Station, Tropical Biosphere Research Center, University of the Ryukyus for their help in collection and maintenance of the coral, M. Hidaka and V. Weis for suggestions, Y. Yokoyama, T. Usami and other staff members of our Unit for sequencing, and S. Brenner, R. Baughman and T. Ichikawa for their support. D. Rokhsar and J. Chapman are acknowledged for suggestions on sequence assembly and gene prediction. The super-computing was supported by the IT Section of OIST and the Human Genome Center, University of Tokyo. This work was supported in part by Grants-in-Aids from MEXT and JSPS, Japan.

**Author Contributions** Project design and coordination: N.S., C.S., E.S., T.K., M.H. and D.J.M. Genome and EST sequencing and assembly: C.S., T.K., E.S., K.H., M. Fujie, M. Fujiwara, M.T., M.H., A.F. and R.K. Annotation and analysis: E.S., C.S., M.H., T.K., K.H., M.T., R.K. and T.I. Writing: N.S., C.S., E.S., T.K., M.H. and D.J.M.

**Author Information** The genome assembly has been deposited with the DNA Databank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL) and GenBank under project accession BACK01000001–BACK01053640 (contigs) and DF093604–DF097774 (scaffolds). Reprints and permissions information is available at www.nature.com/reprints. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike license, and is freely available to all readers at www.nature.com/nature. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to N.S. (norisky@oist.jp).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Biological specimen.** Under permits from the Aquaculture Agency of Okinawa Prefecture (the number 20–27), part of an *A. digitifera* colony was collected and has subsequently been maintained in an aquarium at the Sesoko Station, Tropical Biosphere Research Center, University of the Ryukyus.

**The number of chromosomes, diploidy and genome size of *Acropora digitifera*.** The number of chromosomes was determined by their preparation from nuclei of embryonic cells. The diploidy of the genome was examined by fluorescent *in situ* hybridization (FISH) of BAC clones[31], which were constructed in pKS146 (ref. 32). The genome size was estimated by flow cytometry[33] using sperm nuclei from the same colony that was used to sequence the genome.

**Genome sequencing and assembly.** The sperm was obtained from the single colony and sperm DNA was used for genome sequencing and BAC library construction. Genome sequence data were obtained using single read, paired-end and mate-pair protocols on the Roche 454 GS-FLX[10] and Illumina GAIIx[11] instruments. The genomic DNA was fragmented, libraries prepared and sequencing conducted according to the manufacturer's protocols. The 454 shotgun and paired-end reads were assembled *de novo* by GS *De novo* Assembler version 2.3 (Newbler, Roche)[10] in heterozygotic mode with adjusted algorisms to reflect an increase in the expected variability in sequence identity. Possible PCR duplicates in Illumina mate-pair reads were removed by MarkDuplicates in Picard tools (http://picard.sourceforge.net), and then subsequent scaffolding of the 29,765 Newbler output was performed by SOPRA[27] and SSPACE[28] using the Illumina mate-pair information. Gaps inside the scaffolds were closed with Illumina paired-end data using GapCloser[34]. To overcome potential assembly errors arising from tandem repeats, sequences that were aligned to another sequence over 50% of the length by BLASTN ($1 \times 10^{-50}$) were removed from the assembly[35].

**Transcriptome analyses.** RNA was isolated from eggs, gastrulae, planulae, polyps and adults. Total RNA was extracted following the manufacturer's instructions (Invitrogen) and purified using DNase and an RNeasy micro kit (QIAGEN). Transcriptome libraries for 454 GS-FLX were prepared[36] and sequenced as per manufacturer's instructions. In addition, Illumina 50-bp paired-end RNA-seq sequencing was performed. All high-quality sequences (quality value $\geq$15) were assembled by a Velvet/Oases assembler[37] with hash length 27.

**Gene prediction.** A set of gene model predictions (the *A. digitifera* Gene Model v. 1) was generated using AUGUSTUS[29]. AUGUSTUS 2.0.4 was trained on the 877 EST assemblies recommended by PASA[38] for this purpose. The gene models were created by running AUGUSTUS on a repeat-masked genome produced by RepeatMasker[39], and improved by PASA[38]. A genome browser has been established using the assembled genome sequences using the Generic Genome Browser (GBrowser) 2.17 (ref. 40).

**Identification of *Acropora* genes involved in the response to environmental change.** Three approaches, individual methods or combinations of the methods, were used to annotate the protein-coding genes in the *A. digitifera* genome. A primary approach to the identification of putative orthologues of *A. digitifera* genes was reciprocal BLAST analysis. This was carried out on the basis of mutual best hit in BLAST analyses for human, mouse, or *Drosophila* genes against the *A. digitifera* gene models (BLASTP) or the assembly (BLASTN). A second approach used in the case of genes encoding proteins with one or more specific protein domains, was to screen the merged models against the Pfam database (Pfam-A.hmm, release 24.0; http://pfam.sanger.ac.uk)[30], which contains 11,912 conserved domains using HMMER (hmmer3)[41]. In the case of complex multigene families, a third annotation method was used; sets of related sequences were subjected to phylogenetic analyses to determine more precisely orthology relationships. For these purposes, amino acid sequences were aligned using ClustalW[42] or ClustalX[42] under the default options. Gaps and ambiguous areas were excluded using Gblocks 0.91b[43] with the default parameters and then checked manually. On the basis of the alignment data sets, phylogenetic trees were constructed by neighbour joining and/or maximum likelihood. Calculations and tree construction were performed in SeaView[44]. The KEGG pathway database[14] was used to examine the metabolic repertoire of *Acropora* in comparison to that of the sea anemone *Nematostella*.


31. Shoguchi, E. *et al.* Chromosomal mapping of 170 BAC clones in the ascidian *Ciona intestinalis*. *Genome Res.* **16,** 297–303 (2006).
32. Fujiyama, A. *et al.* Construction and analysis of a human-chimpanzee comparative clone map. *Science* **295,** 131–134 (2002).
33. Davies, D. & Allen, P. in *Flow Cytometry: Principles and Applications* (ed. Macey, M. G.) (Humana, 2007).
34. Li, R. *et al.* The sequence and *de novo* assembly of the giant panda genome. *Nature* **463,** 311–317 (2010).
35. Wurm, Y. *et al.* The genome of the fire ant *Solenopsis invicta*. *Proc. Natl Acad. Sci. USA* **108,** 5679–5684 (2011).
36. Meyer, E. *et al.* Sequencing and *de novo* analysis of a coral larval transcriptome using 454 GSFlx. *BMC Genomics* **10,** 219 (2009).
37. Zerbino, D. R. & Birney, E. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* **18,** 821–829 (2008).
38. Haas, B. J. *et al.* Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **31,** 5654–5666 (2003).
39. Smit, A. F. A., Hubley, R. & Green, P. RepeatMasker Open-3.0. ⟨http://www.repeatmasker.org⟩ (1996–2010).
40. Stein, L. D. *et al.* The generic genome browser: a building block for a model organism system database. *Genome Res.* **12,** 1599–1610 (2002).
41. Eddy, S. R. Profile hidden Markov models. *Bioinformatics* **14,** 755–763 (1998).
42. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948 (2007).
43. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17,** 540–552 (2000).
44. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27,** 221–224 (2010).


©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10276

# Fossil jawless fish from China foreshadows early jawed vertebrate anatomy

Zhikun Gai[1,2], Philip C. J. Donoghue[1], Min Zhu[2], Philippe Janvier[3] & Marco Stampanoni[4,5]

**Most living vertebrates are jawed vertebrates (gnathostomes), and the living jawless vertebrates (cyclostomes), hagfishes and lampreys, provide scarce information about the profound reorganization of the vertebrate skull during the evolutionary origin of jaws[1–9]. The extinct bony jawless vertebrates, or 'ostracoderms', are regarded as precursors of jawed vertebrates and provide insight into this formative episode in vertebrate evolution[8–14]. Here, using synchrotron radiation X-ray tomography[15,16], we describe the cranial anatomy of galeaspids, a 435–370-million-year-old 'ostracoderm' group from China and Vietnam[17]. The paired nasal sacs of galeaspids are located anterolaterally in the braincase, and the hypophyseal duct opens anteriorly towards the oral cavity. These three structures (the paired nasal sacs and the hypophyseal duct) were thus already independent of each other, like in gnathostomes and unlike in cyclostomes and osteostracans (another 'ostracoderm' group), and therefore have the condition that current developmental models regard as prerequisites for the development of jaws[1–3]. This indicates that the reorganization of vertebrate cranial anatomy was not driven deterministically by the evolutionary origin of jaws but occurred stepwise, ultimately allowing the rostral growth of ectomesenchyme that now characterizes gnathostome head development[1–3].**

A key distinction in the structure of the head of cyclostomes, in comparison with gnathostomes, is the presence of a single median duct, or 'nostril', leading to the nasohypophyseal organ, which develops from a single median nasohypophyseal placode[2,3]. The separation of the nasal and hypophyseal placodes and the development of paired, laterally located nasal sacs have been considered fundamental prerequisites for the origin of jaws, because the median position of the nasohypophyseal placode in cyclostome head development precludes the anterolateral growth of the mandibular neural crest from which a major portion of the gnathostome jaw develops[1–3].

The 420–370-million-year-old osteostracans have been regarded as the closest jawless relatives of the gnathostomes[10–13] because they share with the latter a number of uniquely derived characters, such as girdle-supported paired fins, cellular perichondral bone, a sclerotic ring and an epicercal tail[8,9]. However, their massive, mineralized braincase has a median nasohypophyseal organ resembling the condition in lampreys[8,9,18]. Consequently, osteostracans provide little insight into the evolutionary origin of jaws and speculative hypotheses of saltational evolution have therefore been proposed[19]. By contrast, the cranial anatomy of the approximately coeval galeaspids, another 'ostracoderm' group and sister relatives of osteostracans plus gnathostomes[10–13], is poorly known, despite a comparable quality of skeletal preservation[20–22].

Class Galeaspida Tarlo, 1967
Order Eugaleaspiformes Liu, 1980
*Shuyu* gen. nov.

**Etymology.** *Shu* (Chinese Pinyin): dawn; *yu* (Chinese Pinyin): fish. But essentially 'jawed fish'; thus, *Shuyu* here means 'dawn of gnathostomes'. The new genus is erected for *'Sinogaleaspis' zhejiangensis*[23,24] Pan, 1986, from the Maoshan Formation (late Llandovery epoch to early Wenlock epoch, Silurian period, ~430 million years ago) of Zhejiang, China.

**Diagnosis.** Small galeaspid (Supplementary Figs 6 and 7) distinct from *Sinogaleaspis* in its terminally positioned nostril, posterior supraorbital sensory canals not converging posteriorly, median dorsal sensory canals absent, only one median transverse sensory canal and six pairs of lateral transverse sensory canals[17,23].

**Description of cranial anatomy.** To elucidate the gross cranial anatomy of galeaspids, we used synchrotron radiation X-ray tomographic microscopy[15,16] to reconstruct seven skulls of *Shuyu*, one of the earliest and most primitive galeaspid genera. Our interpretation of the endocast is constrained by the comparative anatomy of extant vertebrates that provide a phylogenetic bracket, with hagfishes and lampreys representing a sister clade to 'ostracoderms' plus living jawed vertebrates, and with galeaspids resolved as the sister group of osteostracans plus gnathostomes[10–13,25]. The braincase of *Shuyu* is a massive endoskeletal structure that encloses the brain and sensory organs and wraps a large oralobranchial chamber ventrally (Fig. 1a, b, e, g). The cranial anatomy is inferred from the natural endocast (Fig. 1a, b) or the spaces in the endoskeleton lined post mortem with diagenetic minerals (Fig. 2a). The endocasts preserve the morphology of the brain, nerves, sensory capsules and blood vessels, which we reveal here in three dimensions. Our results confirm features of the galeaspid braincase previously reconstructed on the basis of tenuous evidence[20–22]. They also confirm the remarkable resemblance between galeaspid and osteostracan braincases, except for the arrangement of the forebrain and nasohypophyseal complex.

The galeaspids are characterized by a large median dorsal opening ('no'; Figs 1c–g and 2c–f) in the anterior part of the headshield that serves as both a common nostril and the main water intake device. Beneath the median dorsal opening is a large median oronasal cavity ('on.c'; Supplementary Figs 5 and 7) that also opens ventrally to the mouth ('m'; Fig. 3) and caudally towards the pharynx ('pha'; Fig. 3). The olfactory bulbs are housed in a pair of small fossae ('olf.b'; Figs 1c, d, g and 2a, c–f, h) that are located high in the posterior wall of the oronasal cavity and remain open anteriorly ('pr.f'; Fig. 2a–f, h). Posteriorly, the olfactory bulbs are connected to the cerebral hemispheres ('cer.h'; Fig. 2c) by paired canals for the olfactory tracts ('olf.t'; Figs 1c–e and 2c–e, h), accompanied dorsally by canals for a possible terminal nerve ('ter'; Figs 1e and 2b, h). Laterally, each olfactory bulb receives two canals from the ophthalmic artery, the anterior and posterior ethmoidal arteries ('ae.a' and 'pe.a'; Figs 1d and 2h), which are coupled with two branches of the profundus nerve ('$V_1$'; Figs 1c–e and 2c–e, h). Medially, the paired olfactory bulbs are separated and floored partly by an ethmoid rod ('et.r'; Figs 1d, e, g and 2a, b, h) protruding from what may correspond to the orbitonasal lamina in jawed vertebrates ('on.l'; Supplementary Fig. 5).

The nasal sacs ('na'; Figs 1c–e, g and 2b, g, h) are housed in large, paired, cone-shaped recesses located laterally in the oronasal cavity.

[1]School of Earth Sciences, University of Bristol, Bristol BS8 1RJ, UK. [2]Laboratory of Evolutionary Systematics of Vertebrates, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, China. [3]Muséum National d'Histoire Naturelle, UMR 7207 du CNRS, 47 rue Cuvier, 75231 Paris cedex 05, France. [4]Swiss Light Source, Paul Scherrer Institut, CH-5232 Villigen, Switzerland. [5]Institute for Biomedical Engineering, University and ETH Zurich, 8092 Zurich, Switzerland.



©2011 Macmillan Publishers Limited. All rights reserved

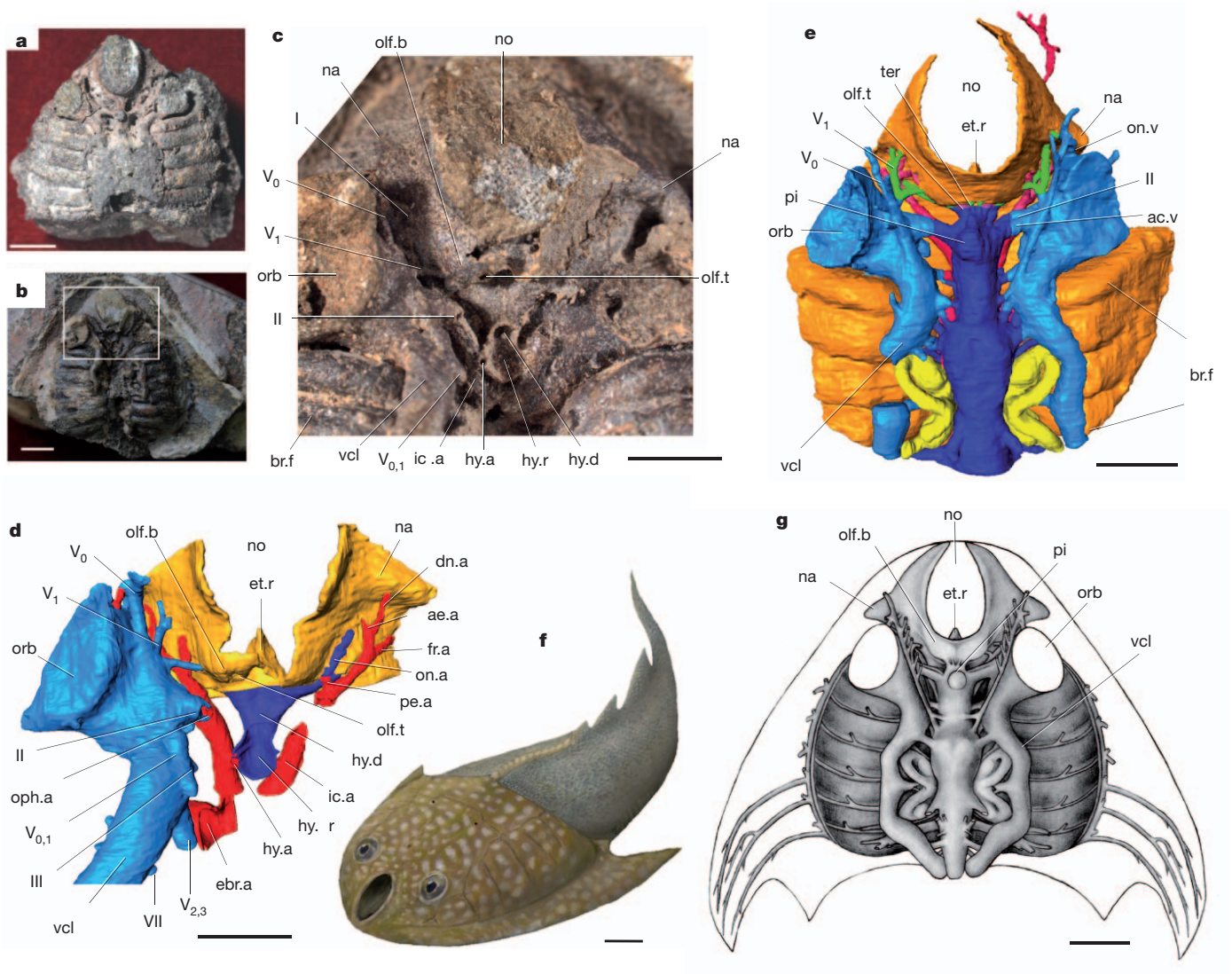



**Figure 1 | *Shuyu zhejiangensis*, Silurian of Zhejiang, China.** **a**, **b**, Two natural endocast specimens: V14334.1 (**a**); V14334.5 (**b**). **c**, **d**, Structure of nasal and hypophyseal region magnified from the boxed region of V14334.5: digital picture (**c**); three-dimensional (3D) reconstruction (**d**). **e**, Virtual endocast (V14334.3). **f**, Restoration of external morphology. **g**, Synthetic restoration. ac.v, anterior cerebral vein; ae.a, anterior ethmoidal artery; br.f, branchial fossa; dn.a, dorsal nasal artery; ebr.a, efferent branchial artery; et.r, ethmoid rod; fr.a, frontal artery; hy.a, hypophyseal artery; hy.d, hypophyseal or buccohypophyseal duct; hy.r, hypophyseal recess; ic.a, internal carotid artery; na, nasal sacs; no, nostril; olf.b, olfactory bulb; olf.t, olfactory tract; on.a, orbitonasal artery; on.v, orbitonasal vein; oph.a, ophthalmic artery; orb, orbital opening; pe.a, posterior ethmoidal artery; pi, pineal organ; ter, terminal nerve; vcl, lateral head vein or dorsal jugular vein; I, II, III, $V_0$, $V_1$, $V_{0,1}$, $V_{2,3}$, VII, olfactory (I), optic (II), oculomotor (III), superficial ophthalmic ($V_0$), profundus ($V_1$), superficial ophthalmic plus profundus ($V_{0,1}$), maxillomandibular ($V_{2,3}$) of trigeminal (V), and facial (VII) nerves. Scale bars, 2 mm.

Posteriorly, each nasal sac is connected to the olfactory bulb by an elongate space for the olfactory nerve ('I'; Fig. 1c) and receives two arterial canals: a dorsal nasal artery ('dn.a'; Figs 1d and 2h) branching off from the ophthalmic artery, and a ventral orbitonasal artery ('on.a'; Fig. 1d), a collateral of the internal carotid artery. Dorsally, the nasal sac is pierced by another canal for the orbitonasal vein ('on.v'; Figs 1e and 2g) that connects posteriorly to the large dorsal jugular vein ('vcl'; Fig. 1c–e, g).

The floor of the diencephalic cavity extends ventrally to form a large funnel-shaped recess for the hypophysis ('hy.r'; Figs 1c, d and 2a, b, d). The hypophyseal recess extends rostrally as a single median horizontal duct for the hypophyseal duct ('hy.d'; Figs 1c, d and 2a, b, d), which opens anteriorly to the oral cavity ('or.c'; Fig. 2a) via the hypophyseal opening ('hy.o'; Fig. 2a, b, d, f, h). The hypophyseal recess is blind caudally (Fig. 2a, b) and shows no posterior communication with the pharynx. It is flanked by a pair of large canals for the internal carotid artery ('ic.a'; Figs 1c, d and 2a, b, d, e). Laterally, it receives a minute canal for the hypophyseal artery ('hy.a'; Figs 1c, d and 2d).

Most strikingly, the paired nasal sacs of galeaspids are located laterally in the anterior part of the braincase, and the hypophyseal duct opens anteriorly and separately towards the oronasal cavity. Thus, in galeaspids these three structures are independent of each other, as in jawed vertebrates, rather than collocated, like the median nasohypophyseal complex of cyclostomes (hagfishes and lampreys) and osteostracans. Distinct, paired nasal sacs are also apparent in the extinct heterostracans[26], but the position of their hypophyseal duct is unknown. The nasohypophyseal complexes of galeaspids and hagfishes have been compared[7,8] but the arrangement of the hagfish hypophyseal duct is quite distinct, opening to the exterior via a prenasal sinus and to the pharynx caudally[27] (Fig. 3). Although the endoskeletal hypophyseal duct of galeaspids communicates with the diencephalic division of the brain cavity, the membranous duct it housed was closed



©2011 Macmillan Publishers Limited. All rights reserved

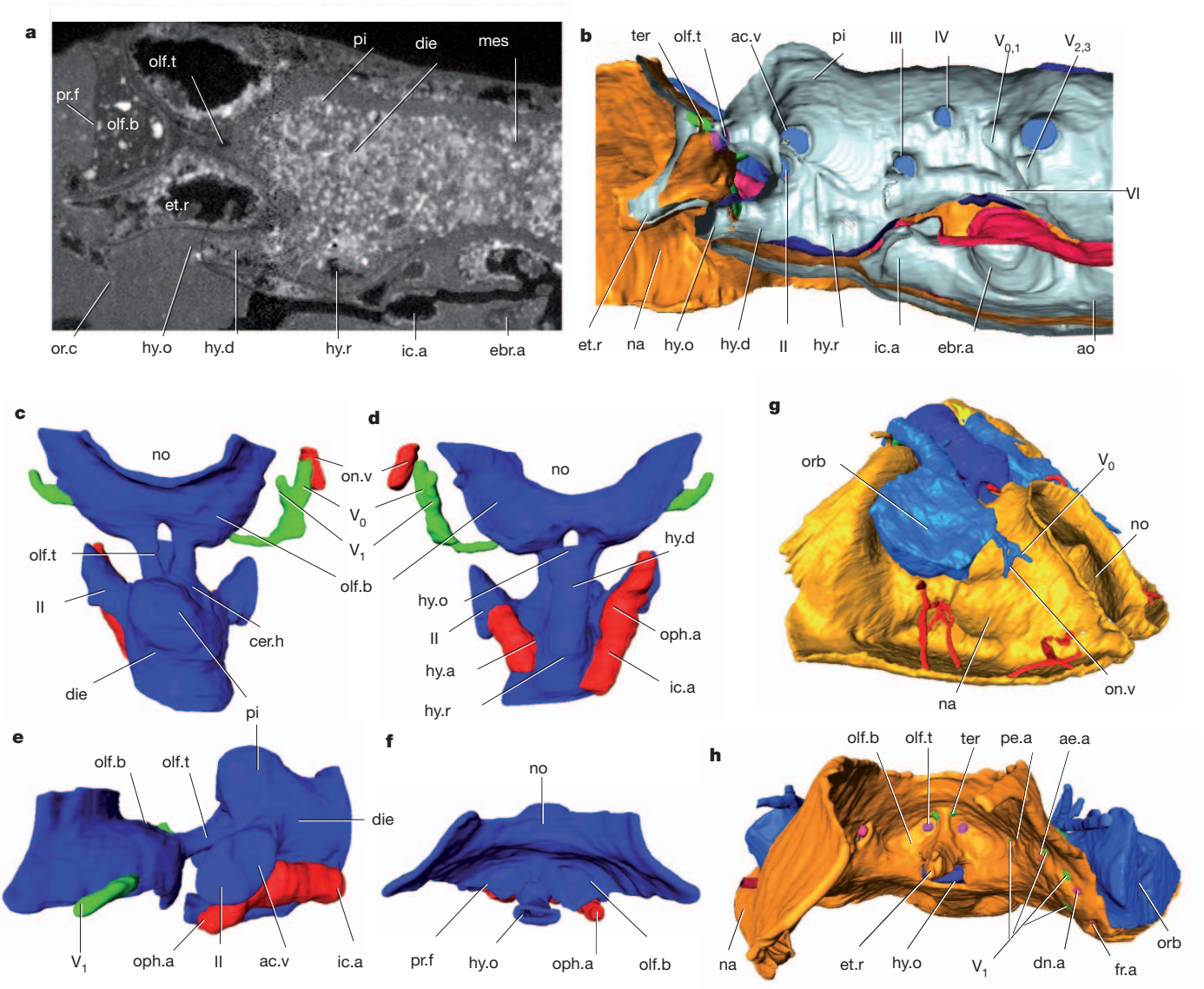



**Figure 2 | The nasohypophyseal complex of *Shuyu zhejiangensis*.** **a**, 512th of 1,024 sagittal slices of V14334.3, from a computerized tomography scan. **b**, Sagittal section through the braincase (V14334.3), showing the relationship between the hypophyseal duct and the nasal sacs. **c**–**f**, 3D reconstruction of the forebrain, hypophysis and related nerves and blood vessels (V14334.7): dorsal view (**c**); ventral view (**d**); lateral view (**e**); frontal view (**f**). **g**, **h**, 3D reconstruction of the nasal region: right oblique lateral view (V14334.2) (**g**); frontal view (V14334.3), showing the clear separation of the nasal sacs and the hypophyseal duct (**h**). ao, dorsal aorta; cer.h, cerebral hemisphere; die, diencephalon; hy.o, hypophyseal or buccohypophyseal opening; mes, mesencephalon; or.c, oral cavity; pr.f, precerebral fenestra; IV, VI, trochlear (IV) and abducens (VI) nerves. See Fig. 1 for other notation.

caudally, as in lampreys, osteostracans and jawed vertebrates (Fig. 3), and could not have conveyed water towards the pharynx. The hypophyseal duct of galeaspids does not show the dorsal bend seen in lampreys and osteostracans, nor the ventral orientation of the buccohypophyseal duct as in jawed vertebrates, perhaps illustrating the intermediate state from which both conditions arose.

The evolution of the nasohypophyseal complex has proven controversial because of the difficulty of distinguishing homologies from parallelisms[7,8]. However, we can now conclude that the arrangement seen in galeaspids, that is, paired nasal sacs flanking an oral cavity into which the hypophyseal duct also opens, is the shared primitive condition inherited by jawed vertebrates. This is because, at the least, separation of paired nasal sacs that seem to open towards the buccal cavity can be seen in heterostracans, the sister lineage to the clade including galeaspids, osteostracans and jawed vertebrates[10–13]. Given this framework of phylogenetic relations, the resemblance between the organization of the nasohypophyseal complex in osteostracans and that in lampreys must be a consequence of convergent evolution (Fig. 3). Osteostracans are derived, like galeaspids, from a generalized stem-gnathostome condition.

Galeaspids provide the earliest evidence for the clear separation of the olfactory organs from the hypophyseal duct in vertebrate phylogeny, a prerequisite condition in evolutionary developmental biology models for the origin of complete diplorhiny[7,8] and jaws[1–3]. Galeaspids reflect an intermediate condition in the establishment of diplorhiny and jaws in which this barrier to the forward growth of neural-crest-derived craniofacial ectomesenchyme was removed. Complete diplorhiny (nasal sacs with individual externalized nostrils) and jaws arose later in the lineage leading to the extinct placoderms and crown-group gnathostomes, but galeaspids foreshadow the stepwise assembly of these characters[8,14], in the absence of any evidence for increasingly active feeding strategies that are commonly invoked to explain the origin of the jawed vertebrates[28,29].



©2011 Macmillan Publishers Limited. All rights reserved

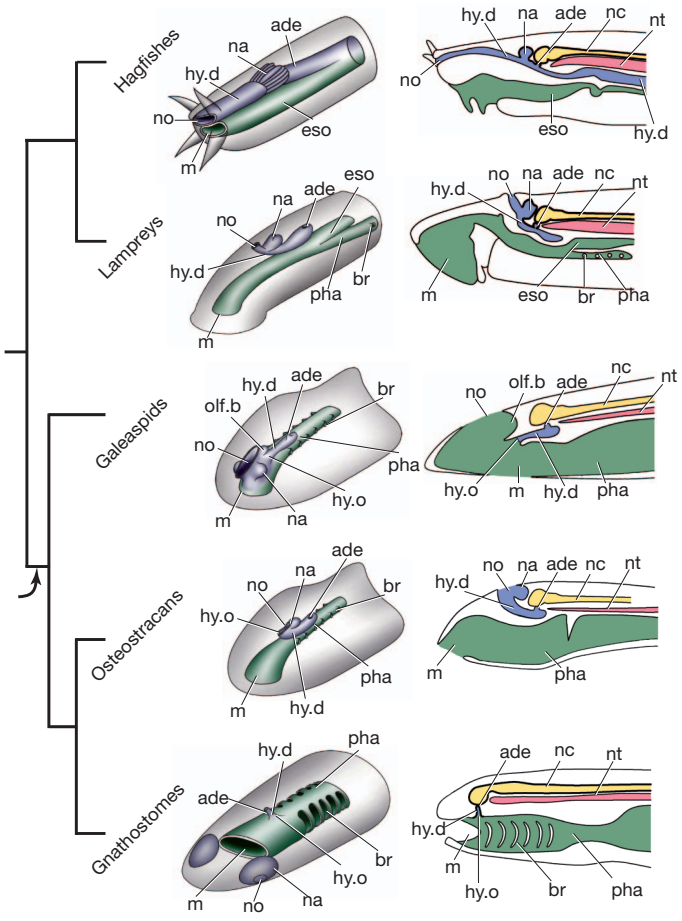



**Figure 3 | The nasohypophyseal complex in craniates.** Left, oblique view; right, sagittal section. The disassociation of the nasohypophyseal complex, an evolutionary prerequisite for the origin of jaws, happened at least in the common ancestor of galeaspids, osteostracans and gnathostomes (arrow). The condition of osteostracans probably converged with that of lampreys. ade, adenohypophysis; br, branchial duct or slit; eso, oesophagus; m, mouth; nc, neural cord; nt, notochord; pha, pharynx. See Figs 1 and 2 for other notation.

## METHODS SUMMARY

Our investigations were performed at the TOMCAT beamline[30] of the Swiss Light Source at the Paul Scherrer Institute. The X-ray energy was optimized for maximum absorption contrast and the magnification of the X-ray microscope was ×4. Isotropic voxels of 1.85 microns were used. Projections (1,501) were acquired equiangularly over 180°, post-processed online and rearranged into flat-field- and dark-field-corrected sinograms. Reconstruction was performed on a 20-node Linux computer cluster using highly optimized filtered back-projection routines. The full volumetric information was available within a few minutes of the end of the scan. Slice data derived from the scans were then analysed and manipulated using AVIZO software (http://www.tgs.com) for computed tomography on a Hewlett Packard Workstation with a 2-GHz Intel processor and 16 GB of random-access memory. Specimens are housed in the Institute of Vertebrate Paleontology and Paleoanthropology, Beijing.

**Received 16 February; accepted 7 June 2011.**


1. Kuratani, S. & Ota, K. G. Primitive versus derived traits in the developmental program of the vertebrate head: views from cyclostome developmental studies. *J. Exp. Zool.* **310B,** 294–314 (2008).
2. Kuratani, S., Nobusada, Y., Horigome, N. & Shigetani, Y. Embryology of the lamprey and evolution of the vertebrate jaw: insights from molecular and developmental perspectives. *Phil. Trans. R. Soc. B* **356,** 1615–1632 (2001).
3. Uchida, K., Murakami, Y., Kuraku, S., Hirano, S. & Kuratani, S. Development of the adenohypophysis in the lamprey: evolution of epigenetic patterning programs in organogenesis. *J. Exp. Zool.* **300B,** 32–47 (2003).
4. Shigetani, Y. *et al.* Heterotopic shift of epithelial-mesenchymal interactions in vertebrate jaw evolution. *Science* **296,** 1316–1319 (2002).
5. Khonsari, R. H., Li, B., Vernier, P., Northcutt, R. G. & Janvier, P. Agnathan brain anatomy and craniate phylogeny. *Acta Zool.* **90,** 52–68 (2009).
6. Smeets, W. J. A. in *The Central Nervous System of Vertebrates* (eds Nieuwenhuys, R., ten Donlelaar, H. J. & Nicholson, C.) 551–654 (Springer, 1998).
7. Janvier, P. *Early Vertebrates* (Clarendon, 1996).
8. Janvier, P. in *Major Events in Early Vertebrate Evolution* (ed. Ahlberg, P. E.) 172–186 (Taylor and Francis, 2001).
9. Janvier, P. in *Major Transitions in Vertebrate Evolution* (eds Anderson, J. S. & Sues, H.-D.) 57–121 (Indiana Univ. Press, 2007).
10. Janvier, P. The dawn of the vertebrates: characters versus common ascent in the rise of current vertebrate phylogenies. *Palaeontology* **39,** 259–287 (1996).
11. Forey, P. L. & Janvier, P. Agnathans and the origin of jawed vertebrates. *Nature* **361,** 129–134 (1993).
12. Forey, P. L. Agnathans recent and fossil, and the origin of jawed vertebrates. *Rev. Fish Biol. Fish.* **5,** 267–303 (1995).
13. Donoghue, P. C. J. & Smith, M. P. The anatomy of *Turinia pagei* (Powrie) and the phylogenetic status of the Thelodonti. *Trans. R. Soc. Edinb. Earth Sci.* **92,** 15–37 (2001).
14. Donoghue, P. C. J. & Purnell, M. A. Genome duplication, extinction and vertebrate evolution. *Trends Ecol. Evol.* **20,** 312–319 (2005).
15. Donoghue, P. C. J. *et al.* Synchrotron X-ray tomographic microscopy of fossil embryos. *Nature* **442,** 680–683 (2006).
16. Tafforeau, P. *et al.* Applications of X-ray synchrotron microtomography for non-destructive 3D studies of paleontological specimens. *Appl. Phys. A* **83,** 195–202 (2006).
17. Zhu, M. & Gai, Z.-K. Phylogenetic relationships of galeaspids (Agnatha). *Vertebr. PalAsiat.* **44,** 1–27 (2006).
18. Stensiö, E. A. *The Downtonian and Devonian Vertebrates of Spitsbergen. Part 1: Family Cephalaspidae.* (Arno, 1927).
19. Mazan, S., Jaillard, D., Baratte, B. & Janvier, P. *Otx1* gene-controlled morphogenesis of the horizontal semicircular canal and the origin of the gnathostome characteristics. *Evol. Dev.* **2,** 186–193 (2000).
20. Halstead, L. B. Internal anatomy of the polybranchiaspids (Agnatha, Galeaspida). *Nature* **282,** 833–836 (1979).
21. Halstead, L. B., Liu, Y.-H., & P'an, K. Agnathans from the Devonian of China. *Nature* **282,** 831–833 (1979).
22. Wang, N.-Z. in *Early Vertebrates and Related Problems in Evolutionary Biology* (eds Chang, M.-M., Liu, Y.-H. & Zhang, G.-R.) 41–65 (Science Press, 1991).
23. Gai, Z.-K., Zhu, M. & Zhao, W.-J. New material of eugaleaspids from the Silurian of Changxing, Zhejiang, China, with a discussion on the eugaleaspid phylogeny. *Vertebr. PalAsiat.* **43,** 61–75 (2005).
24. Pan, J. in *Professional Papers Presented to Professor Yoh Sen-shing* 67–75 (Geological Publishing House, 1986).
25. Heimberg, A. M. Cowper-Sal·lari, R. Sémon, M., Donoghue, P. C. J. & Peterson, K. J. MicroRNAs reveal the interrelationships of hagfish, lampreys, and gnathostomes and the nature of the ancestral vertebrate. *Proc. Natl Acad. Sci. USA* **107,** 19379–19383 (2010).
26. Janvier, P. & Blieck, A. New data on the internal anatomy of the Heterostraci (Agnatha), with general remarks on the phylogeny of the Craniota. *Zool. Scr.* **8,** 287–296 (1979).
27. von Kupffer, C. *Studien zur Vergleichenden Entwicklungsgeschichte des Kopfes der Kranioten. Heft 4: Zur Kopfentwicklung von Bdellostoma* (Lehmann, 1900).
28. Gans, C. & Northcutt, R. G. Neural crest and the origin of the vertebrates: a new head. *Science* **220,** 268–273 (1983).
29. Baker, C. V. H. & Schlosser, G. The evolutionary origin of neural crest and placodes. *J. Exp. Zool.* **304B,** 269–273 (2005).
30. Stampanoni, M. *et al.* TOMCAT: a beamline for tomographic microscopy and coherent radiology experiments. *Synchr. Radiat. Instrum.* **879,** 848–851 (2007).


**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank F. Marone, S. Bengtson, E.-M. Friis, N. J. Gostling, T. Huldtgren, M. Pawlowska and C.-W. Thomas for assistance in retrieving synchrotron radiation X-ray tomographic microscopy data; Q.-S. Chen and W.-J. Zhao for field work; and S. Powell, F.-X. Wu, B. Choo and R.-D. Zhao for illustrations. This work was supported by the Chinese Academy of Sciences (KZCX2-YW-156), the Chinese Foundation of Natural Sciences (40872020, 40930208), the Major Basic Research Projects (2006CB806400) of MST of China, the Paul Scherrer Institut (P.C.J.D.), European Union FP6 (P.C.J.D.), the Leverhulme Trust (P.C.J.D.), the Natural Environmental Research Council (P.C.J.D.), the SYNTHESYS Project (Z.G.) and a Dorothy Hodgkin studentship from the Royal Society (Z.G.).

**Author Contributions** P.C.J.D. and M.Z. designed the project. Z.G. performed the research and led the writing of the manuscript. P.C.J.D., M.S. and Z.G. conducted the synchrotron radiation X-ray tomographic microscopy. All authors discussed the results and wrote the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.C.J.D. (phil.donoghue@bristol.ac.uk) or M.Z. (zhumin@ivpp.ac.cn).



©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10278

# Collaboration encourages equal sharing in children but not in chimpanzees

Katharina Hamann[1], Felix Warneken[2], Julia R. Greenberg[3] & Michael Tomasello[1]

**Humans actively share resources with one another to a much greater degree than do other great apes, and much human sharing is governed by social norms of fairness and equity[1–3]. When in receipt of a windfall of resources, human children begin showing tendencies towards equitable distribution with others at five to seven years of age[4–7]. Arguably, however, the primordial situation for human sharing of resources is that which follows cooperative activities such as collaborative foraging, when several individuals must share the spoils of their joint efforts[8–10]. Here we show that children of around three years of age share with others much more equitably in collaborative activities than they do in either windfall or parallel-work situations. By contrast, one of humans' two nearest primate relatives, chimpanzees (*Pan troglodytes*), 'share' (make food available to another individual) just as often whether they have collaborated with them or not. This species difference raises the possibility that humans' tendency to distribute resources equitably may have its evolutionary roots in the sharing of spoils after collaborative efforts.**

Among great apes, only humans are true collaborative foragers[8,9,11]. Other apes forage in small parties, but they do not actively work together jointly to produce food—the only exception being chimpanzee group-hunting of monkeys[12,13]. In contrast, humans in all societies produce significant portions of their food through collaborative efforts, even bringing the results of their labour back to some central location to share with other group members[14,15]. After group-hunting, chimpanzees mostly share only under pressure of harassment by others[16] or else reciprocally with coalition partners[17].

Human children actively share valuable resources with others to some degree from early in ontogeny. A fairly well-established pattern across cultures is that three- to four-year-old children tend to divide a windfall of resources unequally, keeping the majority for themselves[4–6,18,19]. As they approach school age, they begin to share more equally[4,5,7,18,19]. But given that humans generate many or most of their resources collaboratively, a plausible hypothesis is that children would share a resource more equitably at an earlier age if it was not provided by adults as a windfall, but if instead they had to work together to produce it[20]. Furthermore, we might expect this positive effect of collaboration on sharing to be confined to humans, among great apes, as only they have an evolutionary history of obligate collaborative foraging[8,9,11].

In the current series of experiments, therefore, we presented pairs of human children and pairs of chimpanzees with resource distribution problems in which one individual had control of more than half of the resources and could choose whether or not to share them equally with their partner. The basic variable was whether the initial unequal distribution of resources resulted from a collaborative effort in which each contributed equally, or whether it came from some non-collaborative source (for example as a windfall or as a result of each individual working on their own).

In study 1, pairs of either two- or three-year-old children were in a room by themselves. In the 'collaboration' condition, they faced an enclosed board with a rope extruding from each end (Fig. 1a), and they knew from previous experience (from a demonstration phase) that they had to pull together to bring the board towards them. On each end of the board were two rewards (small toys) that could be accessed once the board had been pulled close enough. As the children pulled, one of the toys rolled to the other end of the board such that one child ended up with three toys and the other ended up with only one. In the control, 'no-work', condition, by contrast, as children entered the room the board with the toys was already at its end-state position, with three toys at one end and one at the other (Fig. 1b). The main result was that the 'lucky' child, who had gained three toys, made one of the toys available to the 'unlucky' partner, who had gained one, restoring equity, more often in the collaboration condition than in the no-work condition ($F(1, 22) = 21.85$ (analysis of variance), $P < 0.001$). The effect was similar for children of both ages (Fig. 2a).

In this experiment, it was possible that from the beginning of the collaboration children viewed the rewards on their end of the board as belonging to them, such that when one reward rolled to the other end it was as if one of their possessions had been taken away (which was not the case in the no-work condition). In study 2, therefore, we presented pairs of two- or three-year-old children initially with four toys bunched together, so that an initial sense of possession was not an issue. In addition, we added a second control condition—the parallel-work condition—with a very similar set-up, in which each child pulled on a separate board with their own separate rope, to account for the fact that the collaboration condition required work whereas the original control condition (no-work) did not (Fig. 1c–e). Thus, if children are attentive to work effort in general and not to collaborative effort in particular, they should share similarly in the parallel-work and collaboration conditions. However, in this study also, the three-year-old lucky child handed over one of the toys to the unlucky partner more often in the collaboration condition than in either of the two control conditions (no-work and parallel-work). By contrast, the two-year-olds did not differentiate among conditions (see Fig. 2b for data and statistics for both ages).

Because studies 1 and 2 consisted of multiple trials, they leave open the possibility that children shared in the collaboration condition out of a concern that if they did not share their partner might not pull their end of the rope in future trials (which was not an issue in the parallel-work and no-work conditions as children obtained rewards on their own.) In study 3, to ensure that children understood that they would play the game only once, in the demonstration phase we showed them the total number of toys available and made it clear that their number decreased over demonstrations. When there was only one set of four toys left, we pointed this out and specifically asked the children whether the game could be played after this last set was gone. Only children who answered that this would be the last time were then given the actual test trial (that is, only their data was used for analysis; see Supplementary Information for details). Replicating our results once more, three-year-old children equalized the distribution of toys more often in the collaboration trials (75%) than in the parallel-work trials (25%; $\chi^2_{(d.f.=1,n=24)} = 6.0$, $P = 0.039$; Fig. 2c). Taken together, these studies show that collaborative work encourages equal sharing in

[1]Max Planck Institute for Evolutionary Anthropology, Department of Developmental and Comparative Psychology, Deutscher Platz 6, 04103 Leipzig, Germany. [2]Harvard University, Department of Psychology, 33 Kirkland St, Cambridge, Massachusetts 02138, USA. [3]Michigan State University, Department of Zoology, East Lansing, Michigan 48824-1115, USA.



©2011 Macmillan Publishers Limited. All rights reserved

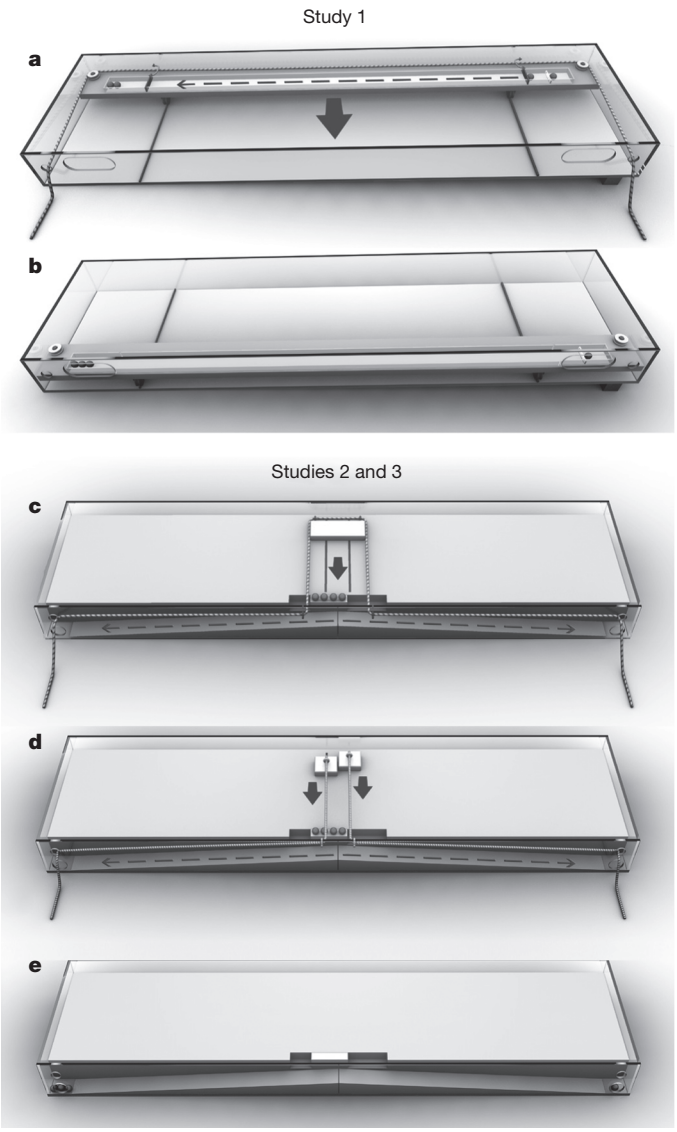



**Figure 1 | Child study tasks.** **a**, Apparatus from study 1 with reward relocation mechanism as used in the collaboration condition (180 cm × 60 cm × 15 cm; adapted from studies with chimpanzees[29,30]). In the collaboration condition, children had to pull both ends of the rope simultaneously to move the board towards the access holes in the front of the enclosure (solid arrow). Initially, two toys (marbles) were on each side (as shown), but as the children pulled the board closer, the black barriers slipped out such that one marble rolled to the other end, resulting in a 3:1 reward distribution (moving from right to left in this example; dashed arrow). **b**, In the no-work condition, the board was already in the front part of the apparatus, with no attached rope, when children approached it (same reward distribution, of 3:1). **c**, In studies 2 and 3, children had to move a block closer to move the marbles such that they would roll in front of the access holes. In the collaboration condition, children had to pull a single, long rope simultaneously to move a large block closer (solid arrow), moving four marbles at once, which then rolled towards the respective access holes (in this example, three marbles rolled to the left and one marble rolled to the right; dashed arrows). **d**, In the parallel-work condition, two smaller blocks (each with a rope attached) could be pulled individually, one by each child, causing the respective marbles to move and roll down the ramps. **e**, The no-work condition, without any work but with the same reward distribution, 3:1.

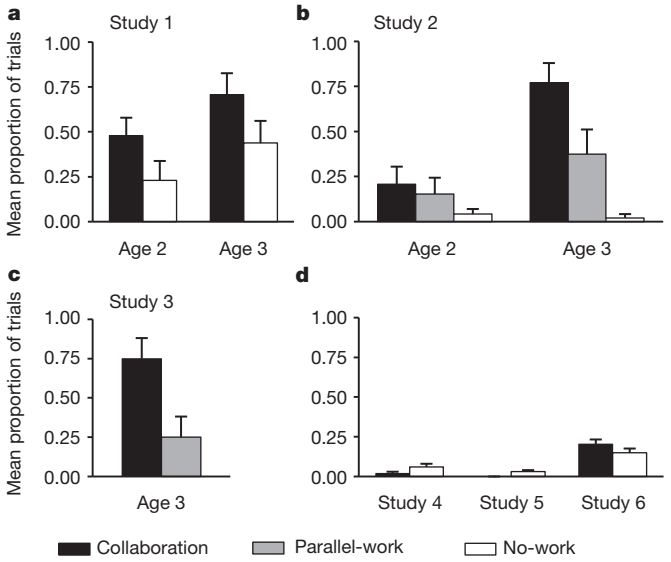



**Figure 2 | Rates of equal shares.** **a**, In study 1, children in both age groups shared more often in the collaboration condition than in the no-work condition ($F(1, 22) = 21.85$, $P < 0.001$). This was true even if only the results of the first trial in both conditions were used for the analysis (McNemar test, $n_{1,0} = 0$ (number of dyads sharing in the no-work condition but not in the collaboration condition), $n_{0,1} = 6$ (number of dyads sharing in the collaboration condition but not in the no-work condition), $P = 0.03$). **b**, In study 2, three-year-olds, but not two-year-olds, shared differently in the three conditions (significant age × condition interaction, $F(2, 66) = 5.26$, $P = 0.008$; main effect of condition, $F(2, 66) = 12.87$, $P < 0.001$). Three-year-olds shared significantly more often in the collaboration condition than in either of the other two conditions (*post hoc* Scheffé tests, both $P < 0.05$). The difference between the parallel-work and no-work conditions approached significance ($P = 0.06$). **c**, In study 3, children shared significantly more often in the collaboration condition as compared with the parallel-work condition ($\chi^2_{(\text{d.f.}=1, n=24)} = 6.0$, $P = 0.039$). **d**, Across studies 4–6, chimpanzees did not share differently in the collaboration and control (no-work) conditions. See main text and Supplementary Information for details and additional analyses. Error bars, s.e.m.

children much more than does working in parallel or acquiring resources in a windfall.

Chimpanzees do not regularly offer resources to others actively, so to test for the same effect of collaboration on sharing in chimpanzees we had to use a slightly more complex apparatus that enabled one individual to provide another with food that the second could not obtain. Although some researchers have proposed that chimpanzees use work effort during group hunts as a criterion for dividing up the spoils[21], our hypothesis was that because chimpanzees are not true collaborative foragers (at least not to the degree of humans[8,9,13]), they would not share differently in windfall and collaboration situations.

The two chimpanzees operated a single apparatus but from adjacent rooms (after enough practice with the apparatus for both to know how it worked; Supplementary Information). The upper level of the apparatus (Fig. 3) was similar to the apparatus used in the experiment with children, in that it contained a long board holding rewards (in this case food) and attached to a rope that each chimpanzee could access. In all conditions, at some point there was one piece of food on a lower level on a see-saw device, such that the lucky chimpanzee could tip the see-saw only either towards itself or away from itself and towards its partner. In a series of three experiments, we made it increasingly easy for one chimpanzee to make the fallen piece of food available to its partner. In all three experiments, there was a collaboration condition, in which the pair worked together to pull on the board, and a control (windfall) condition, in which the food got into position without the chimpanzees' joint effort.

The first chimpanzee study (study 4) was very similar to the first child study (as this procedure was most facilitative of sharing for the children). In this study, the lucky chimpanzee (who got two pieces of food and had a chance to get the fallen piece) could take the fallen piece for itself or could restore the 2:2 balance by actively providing the fallen piece to the unlucky partner (who got one piece of food and had a chance to get the fallen piece) or by doing nothing and letting the unlucky partner tip it to itself. What happened most often was that



©2011 Macmillan Publishers Limited. All rights reserved

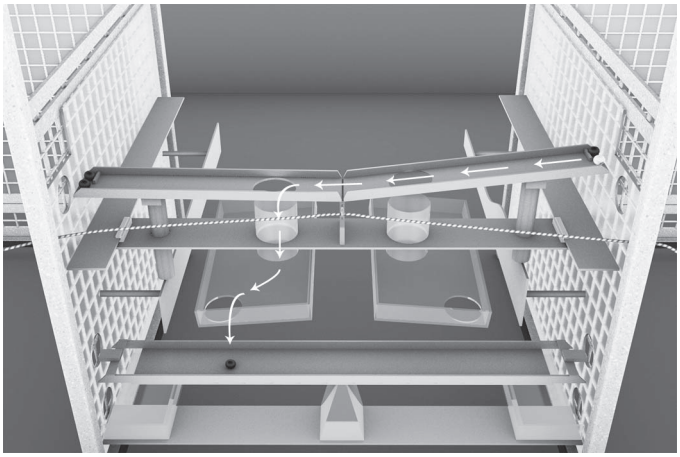


**Figure 3 | Chimpanzee study task.** The apparatus was mounted in a booth between testing rooms. Two chimpanzees located in adjacent rooms (partition mesh not shown) had to pull the two ends of the rope simultaneously to move the upper-level platform in front of the access holes. During the movement, one of the unlucky individual's rewards (grapes) fell onto the lower, see-saw, level and rolled to the lucky chimpanzee's side. In the picture, the sliding platform (upper level) has been pulled almost to the front. This movement has caused one side of the platform to tilt, causing one grape from the right (initial location represented by the grape drawn in white) to roll to the left, falling through a hole and landing on the see-saw mechanism below. The see-saw could then be tilted by the chimpanzees (manner dependent on study and condition).

the unlucky partner almost immediately tipped the see-saw and took the fallen reward for itself (63% of trials); in no cases did the lucky chimpanzee actively tip the reward to the unlucky partner (even though in pre-training they often tipped the food away from themselves into the other room, if they themselves could then go through an open door and get it). In the remaining trials, the lucky chimpanzee took the reward for itself. Importantly, the fate of the fallen reward did not differ between conditions (Wilcoxon signed-rank test, $T+ = 31.5$, $n = 12$ (no ties), $P = 0.60$). Additional analyses showed that in only 4% of cases did the lucky chimpanzee in fact give up the food voluntarily by tolerating the unlucky partner's taking of it (again with no difference between conditions: $T+ = 18$, $n = 11$ (one tie), $P = 0.19$, Fig. 2d).

In the second chimpanzee study (study 5), we tried to encourage the lucky partner by making it impossible for the unlucky partner to operate the see-saw. The result was that the lucky chimpanzee almost always tipped the food to itself (98% of trials), thus creating a 3:1 reward imbalance. There was again no difference in the rate of equitable sharing between the collaboration and control conditions (Wilcoxon signed-rank test, $T+ = 0$, $n = 4$ (eight ties), $P = 0.13$). In the third chimpanzee study (study 6), we made it impossible for the lucky partner to get the fallen piece of food for itself: if it tipped the see-saw towards itself, the food was lost. This new set-up resulted in a higher sharing rate than before (mean, 0.17 of trials; s.d. = 0.17), with subjects tipping the reward to the unlucky partner more often than in both chimpanzee study 1 ($T+ = 69$, $n = 12$ (no ties), $P = 0.016$) and chimpanzee study 2 ($T+ = 64$, $n = 11$ (one tie), $P = 0.003$). However, as in the previous two chimpanzee experiments, the results did not differ between the collaboration and control conditions ($T+ = 40$, $n = 10$ (two ties), $P = 0.22$; Fig. 2d).

Previous research with older school-age children (seven to ten years of age) has shown that they take into account work effort in so-called distributive justice problems in which one individual must say how the fruits of a collaborative effort should be doled out to participants[22–26]. Younger children typically are not able to factor work effort into their decisions in this way. Nevertheless, the current study shows that although they may be unable to balance work and rewards sensitively, children as young as two or three years of age do take note of whether or not rewards were produced from collaborative efforts with others, and that this affects how they think the rewards should be distributed. Thus, the ontogenetically first sense of distributive justice may be that participation in a collaborative effort demands an equal division of spoils. Because chimpanzees rely very little on collaboration for subsistence, they have not evolved the tendency to distribute resources more equally when those resources result from a collaboration.

Collaborative foraging, by definition, requires partners. Any individual with a tendency to take more than their share of the fruits of a collaboration would not be chosen as a partner very often[27,28]. A possible evolutionary picture is thus that this 'social selection' of a tendency to share the fruits of collaboration equally among participants became ever stronger as the need to work together jointly in subsistence activities became ever more obligate. The current results, according to which young children, but not chimpanzees, share more equally after collaboration than in other situations, provide at least indirect support for this picture.

## METHODS SUMMARY

**Children.** We tested children at 2 and 3 years of age (study 1, $n = 48$; study 2, $n = 144$; study 3, $n = 48$) who were paired with a same-sex peer from the same kindergarten. In the demonstration phase of each study, we first familiarized children with the apparatus requiring them to pull ropes to retrieve rewards (marbles to play an individual game). In all studies, the test event was that the lucky child ended up with three marbles and the unlucky child ended up with only one.

In study 1, during the demonstration phase both individuals learned how to pull together to bring an enclosed board holding rewards within reach of access holes in the enclosure (Fig. 1a). In the test phase, we presented two conditions, collaboration and no-work, in counterbalanced order within dyads.

In study 2, different pairs of children worked on a slightly modified version of the original apparatus (Fig. 1c). Children were tested either in a collaboration condition (preceded by a collaborative demonstration as in study 1), a parallel-work condition (preceded by an individual work demonstration; Fig. 1d) or a no-work condition (preceded by a joint no-work demonstration; Fig. 1e).

In study 3, pairs of three-year-old children participated either in a single collaboration test trial or in a single parallel-work test trial (preceded by demonstrations similar to those in study 2.)

**Chimpanzees.** We tested 12 chimpanzees separately with two partners from their social group, for a total of 12 test pairs. We conducted three studies with increasing levels of encouragement of the lucky individual to share. We achieved this by blocking the holes the chimpanzees could use to tip the see-saw or retrieve the food (Fig. 3). In all three experiments, we presented the collaboration and no-work conditions in a within-subject design administered in counterbalanced order (for details, see Supplementary Information).

**Received 12 September 2010; accepted 9 June 2011.**
**Published online 20 July 2011.**

1. Fehr, E. & Fischbacher, U. The nature of human altruism. *Nature* **425,** 785–791 (2003).
2. Gurven, M. To give or not to give: an evolutionary ecology of human food transfers. *Behav. Brain Sci.* **27,** 543–583 (2004).
3. Henrich, J. *et al.* ''Economic man'' in cross-cultural perspective: behavioral experiments in 15 small-scale societies. *Behav. Brain Sci.* **28,** 795–855 (2005).
4. Benenson, J. F., Pascoe, J. & Radmore, N. Children's altruistic behavior in the dictator game. *Evol. Hum. Behav.* **28,** 168–175 (2007).
5. Fehr, E., Bernhard, H. & Rockenbach, B. Egalitarianism in young children. *Nature* **454,** 1079–1083 (2008).
6. Blake, P. R. & Rand, D. G. Currency value moderates equity preference among young children. *Evol. Hum. Behav.* **31,** 210–218 (2010).
7. Gummerum, M., Hanoch, Y., Keller, M., Parsons, K. & Hummel, A. Preschoolers' allocations in the dictator game. The role of moral emotions. *J. Econ. Psychol.* **31,** 25–34 (2010).
8. Tomasello, M. *Why We Cooperate* (MIT Press, 2009).
9. Sterelny, K. Social intelligence, human intelligence and niche construction. *Phil. Trans. R. Soc. B* **362,** 719–730 (2007).
10. de Waal, F. B. M. & Berger, M. L. Payment for labour in monkeys. *Nature* **404,** 563 (2000).
11. Hill, K. Altruistic cooperation during foraging by the Ache, and the evolved human predisposition to cooperate. *Hum. Nature* **13,** 105–128 (2002).
12. Boesch, C. & Boesch, H. Hunting behavior of wild chimpanzees in the Tai-National-Park. *Am. J. Phys. Anthropol.* **78,** 547–573 (1989).
13. Muller, M. & Mitani, J. C. in *Advances in the Study of Behavior* (eds Slater, P. J. B., Rosenblatt, J., Snowdon, C., Roper, T. & Naguib, M.) 275–331 (Elsevier, 2005).



©2011 Macmillan Publishers Limited. All rights reserved

14. Hill, K., Barton, M. & Hurtado, A. M. The emergence of human uniqueness: characters underlying behavioral modernity. *Evol. Anthropol.* **18,** 187–200 (2009).
15. Marlowe, F. W. Hunter-gatherers and human evolution. *Evol. Anthropol.* **14,** 54–67 (2005).
16. Gilby, I. C. Meat sharing among the Gombe chimpanzees: harassment and reciprocal exchange. *Anim. Behav.* **71,** 953–963 (2006).
17. Mitani, J. C. & Watts, D. Why do chimpanzees hunt and share meat? *Anim. Behav.* **61,** 915–924 (2001).
18. Thompson, C., Barresi, J. & Moore, C. The development of future-oriented prudence and altruism in preschoolers. *Cogn. Dev.* **12,** 199–212 (1997).
19. Rochat, P. *et al.* Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *J. Cross Cult. Psychol.* **40,** 416–442 (2009).
20. Warneken, F., Lohse, K., Melis, A. P. & Tomasello, M. Young children share the spoils after collaboration. *Psychol. Sci.* **22,** 267–273 (2011).
21. Boesch, C. Cooperative hunting in wild chimpanzees. *Anim. Behav.* **48,** 653–667 (1994).
22. Lerner, M. J. The justice motive: ''equity'' and ''parity'' among children. *J. Pers. Soc. Psychol.* **29,** 539–550 (1974).
23. Damon, W. *The Social World of the Child* (Jossey-Bass, 1977).
24. Hook, J. G. & Cook, T. D. Equity theory and the cognitive ability of children. *Psychol. Bull.* **86,** 429–445 (1979).
25. Kienbaum, J. & Wilkening, F. Children's and adolescents' intuitive judgments about distributive justice: integrating need, effort, and luck. *Eur. J. Dev. Psychol.* **6,** 481–498 (2009).
26. Almås, I., Cappelen, A. W., Sørensen, E., Ø. & Tungodden, B. Fairness and the development of inequality acceptance. *Science* **328,** 1176–1178 (2010).
27. Boehm, C. Impact of the human egalitarian syndrome on Darwinian selection mechanics. *Am. Nat.* **150,** S100–S121 (1997).
28. Kurzban, R. & Leary, M. R. Evolutionary origins of stigmatization: the functions of social exclusion. *Psychol. Bull.* **127,** 187–208 (2001).
29. Hirata, S. & Fuwa, K. Chimpanzees (*Pan troglodytes*) learn to act with other individuals in a cooperative task. *Primates* **48,** 13–21 (2006).
30. Melis, A. P., Hare, B. & Tomasello, M. Engineering cooperation in chimpanzees: tolerance constraints on cooperation. *Anim. Behav.* **72,** 275–286 (2006).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank all the children and day-care centres involved in this work for their cooperation, and we thank J. Dauksch, G. Duden, A. Gampe, I. Gavriliu, L. Hering, C. Markmann, M. Plötner, E. Rossi, A. Schrimpf, C. Weiske and the practical students at the Wolfgang Köhler Primate Research Center, Leipzig, for their help in recruiting children, data collection and behaviour coding. We also thank R. Pieszek and M. Ulrich for building the apparatuses, S. Tüpke and M. Sureck for creating the figures, R. Mundry for statistical advice and the zookeepers at Leipzig Zoo for their assistance with the chimpanzees. This work received funding from the EC Seventh Framework Program (FP7/2007-2013) under grant agreement no. 215805.

**Author Contributions** K.H., F.W., J.R.G. and M.T. designed the study and wrote the paper. K.H. and J.R.G. conducted the studies and analysed the data. F.W. supervised the research.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.H. (khamann@eva.mpg.de) or F.W. (warneken@wjh.harvard.edu).



©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10272

# 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen

Ken-ichiro Taoka[1]*, Izuru Ohki[2]*, Hiroyuki Tsuji[1]*, Kyoko Furuita[2], Kokoro Hayashi[2], Tomoko Yanase[1], Midori Yamaguchi[1], Chika Nakashima[1], Yekti Asih Purwestri[1]†, Shojiro Tamaki[1], Yuka Ogaki[1], Chihiro Shimada[1], Atsushi Nakagawa[3], Chojiro Kojima[2,3] & Ko Shimamoto[1]

**'Florigen' was proposed 75 years ago[1] to be synthesized in the leaf and transported to the shoot apex, where it induces flowering. Only recently have genetic and biochemical studies established that florigen is encoded by *FLOWERING LOCUS T* (*FT*), a gene that is universally conserved in higher plants[2–4]. Nonetheless, the exact function of florigen during floral induction remains poorly understood and receptors for florigen have not been identified. Here we show that the rice FT homologue Hd3a[5] interacts with 14-3-3 proteins in the apical cells of shoots, yielding a complex that translocates to the nucleus and binds to the *Oryza sativa* (Os)FD1 transcription factor, a rice homologue of *Arabidopsis thaliana* FD. The resultant ternary 'florigen activation complex' (FAC) induces transcription of *OsMADS15*, a homologue of *A. thaliana APETALA1* (*AP1*), which leads to flowering. We have determined the 2.4 Å crystal structure of rice FAC, which provides a mechanistic basis for florigen function in flowering. Our results indicate that 14-3-3 proteins act as intracellular receptors for florigen in shoot apical cells, and offer new approaches to manipulate flowering in various crops and trees.**

In the current model, florigen is transported from leaves to the shoot apex, where it interacts with the basic leucine zipper (bZIP) transcription factor FD to activate expression of floral-meristem identity genes such as *AP1* (refs 6, 7). We previously demonstrated that the rice FT homologue Hd3a is a strong candidate for florigen, based on an analysis of transgenic rice expressing GFP-tagged Hd3a[5]. Furthermore, we obtained evidence that another homologue of Hd3a, RICE FLOWERING LOCUS T1 (RFT1), also acts as a rice florigen under long-day conditions[8], and that Hd3a and RFT1 activate the expression of *OsMADS15* (refs 8–10), indicating that the mechanism of florigen action is conserved in rice and *Arabidopsis*.

To understand more precisely how the Hd3a florigen induces flowering, we performed yeast two-hybrid screening[11] and identified GF14c (a rice 14-3-3 protein), OsKANADI1 (a rice homologue of *Arabidopsis* KANADI) and OsBIP116b (a rice homologue of *Arabidopsis* WAVE-DAMPENED2) as three Hd3a-interacting partners (Fig. 1a). It had previously been reported that 14-3-3 proteins interact with FT proteins[12,13]. We also identified OsFD1, a rice FD homologue (Supplementary Fig. 1). Three Hd3a interactors, but not GF14c, share a common SAP motif (Ser-Ala-Pro) (Fig. 1a), which resembles the FT-interacting motif in FD[6]. Alanine substitutions in the SAP motifs abolished the interactions with Hd3a (Fig. 1b), indicating that the SAP motif is common to FT/Hd3a interactors. We tested for a direct interaction between highly purified Hd3a and OsFD1 using three different methods, but no interaction was detected (Fig. 1c and Supplementary Fig. 2). By contrast, the known interaction between Hd3a and GF14c (ref. 11) was observed (Fig. 1c–d and Supplementary Figs 3 and 4). GF14c also interacted directly with OsFD1 in glutathione *S*-transferase (GST) pull-down assays (Fig. 1c).

On the basis of these experiments, and the observation that the SAP motif is similar to conventional 14-3-3 recognition motifs (R/K-S-X-P and R/K-X-X-S-X-P) and is absent in 14-3-3 proteins, we hypothesized that 14-3-3 proteins might mediate interactions between Hd3a and the identified proteins. In our hypothesis, endogenous yeast 14-3-3 proteins could have bridged Hd3a and the identified Hd3a-interacting proteins in the yeast two-hybrid assays. We examined the interaction between OsFD1 and GF14b and GF14c (Fig. 1e). Mutation analysis of the SAP motif in OsFD1 indicated that the OsFD1–GF14 interaction depends on the phosphorylation of S192. Interaction between Hd3a and 14-3-3 in the shoot apex of rice plants was confirmed by co-immunoprecipitation assays (Fig. 1f).

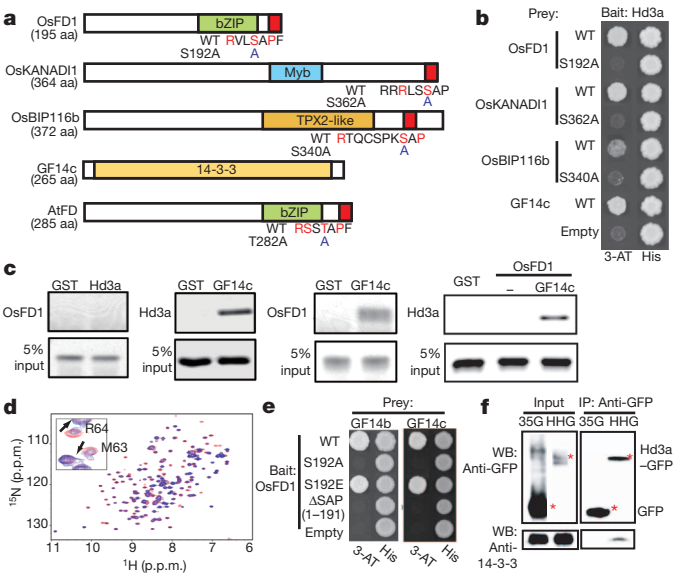



**Figure 1 | Interactions among Hd3a, GF14 and OsFD1. a**, Schematic representation of the Hd3a interactors. Red boxes, SAP motifs; aa, amino acids; WT, wild type; TPX2, targeting protein for Xklp2. **b**, Yeast two-hybrid assay between Hd3a and the identified interactors. 3-AT, 3-aminotriazole; His, histidine. **c**, GST pull-down assays using (from left to right) GST–Hd3a and OsFD1; GST–GF14c and Hd3a; GST–GF14c and OsFD1; GST–OsFD1 and Hd3a with or without GF14c. **d**, Chemical shift perturbation of Hd3a. $^{1}H$–$^{15}N$ heteronuclear single quantum coherence (HSQC) spectra of $^{15}N$-labelled Hd3a were recorded in the presence (blue) and absence (red) of non-labelled GF14c. p.p.m., parts per million. **e**, Yeast two-hybrid assay between OsFD1 mutants and GF14b or GF14c. **f**, Co-immunoprecipitation assay of Hd3a and GF14 in the shoot apex. Transgenic rice plants transformed with *CaMV 35S promoter:GFP* (35G) or *Hd3a promoter:Hd3a–GFP* (HHG) were used. IP, immunoprecipitation; WB, western blot; *, GFP signal.

[1]Laboratory of Plant Molecular Genetics, Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan. [2]Laboratory of Biophysics, Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan. [3]Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan. †Present address: Faculty of Biology, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.
*These authors contributed equally to this work.



©2011 Macmillan Publishers Limited. All rights reserved

To elucidate the structural basis of the interactions between the three proteins, we determined the crystal structure of an Hd3a–GF14c–OsFD1 complex at 2.4 Å resolution ($R_{cryst}$ 22.7% and $R_{free}$ 27.6%) (Fig. 2a, b and Supplementary Table 1). For structure determination, GF14c and a peptide comprising the nine carboxy-terminal amino acids of phosphorylated OsFD1 were used (Supplementary Figs 5 and 6 and Supplementary Text 1). The complex is a hetero-hexamer composed of two molecules each of Hd3a, GF14c and OsFD1. Two Hd3a monomers abut the C-terminal regions of dimeric GF14c, about 50 Å apart, resembling two equal weights on a balance, to form a thick and deep W-shaped structure (Fig. 2a, b). The two corners at the inner base of the 'W' form positively charged pockets (Supplementary Fig. 7), and the C terminus of OsFD1, phosphorylated at S192, binds to these pockets. The binding sites in GF14c for Hd3a are separated by more than 20 Å from those for OsFD1 (Fig. 2a, b), and yeast two-hybrid assays using GF14 mutants confirmed that the two partners bound independently (Supplementary Figs 6 and 8). GST pull-down assays showed that Hd3a and full-length OsFD1 could not interact in the absence of GF14c (Fig. 1c). Therefore, GF14c forms a stable complex with Hd3a and OsFD1 simultaneously, and mediates indirect binding between Hd3a and OsFD1. We named this complex the florigen activation complex (FAC).

The Hd3a binding site in GF14c is far apart from the canonical phosphoserine binding site[14], but is close to the Cdc25C binding site predicted by a mutation study of human 14-3-3σ[15]. The Hd3a binding surface of GF14c (~1,300 Å$^2$) is well determined (Fig. 2c) and consists of two major areas: a hydrophobic cavity composed of helices H and I, and an acidic lobe comprising the loop between these two helices (residues 206–215) (Supplementary Fig. 9 and Supplementary Text 2). This acidic lobe interacts with Hd3a R64 and R132 via salt bridges. The surface of GF14c that binds to the phosphorylated OsFD1 fragment is also well determined (Fig. 2d and Supplementary Fig. 7). The interface topology resembles the canonical mode-I-type interaction between 14-3-3 proteins and phosphorylated peptides[14]: GF14c interacts specifically with a characteristic tight turn created by the phosphoserine and *cis*-proline (residues 192 and 194 in the SAP motif) of phosphorylated OsFD1 (Supplementary Text 3). Amino acids located at the interface were mutated to alanine and other residues, and the consequent binding between Hd3a, OsFD1 and 14-3-3 proteins was assessed by GST pull-down assay (Supplementary Fig. 14), NMR (Fig. 1d and Supplementary Fig. 15) and yeast two-hybrid assay (Supplementary Figs 6 and 8). These experiments confirmed interaction data obtained from structural analyses of the FAC (Supplementary Text 2 and 3).

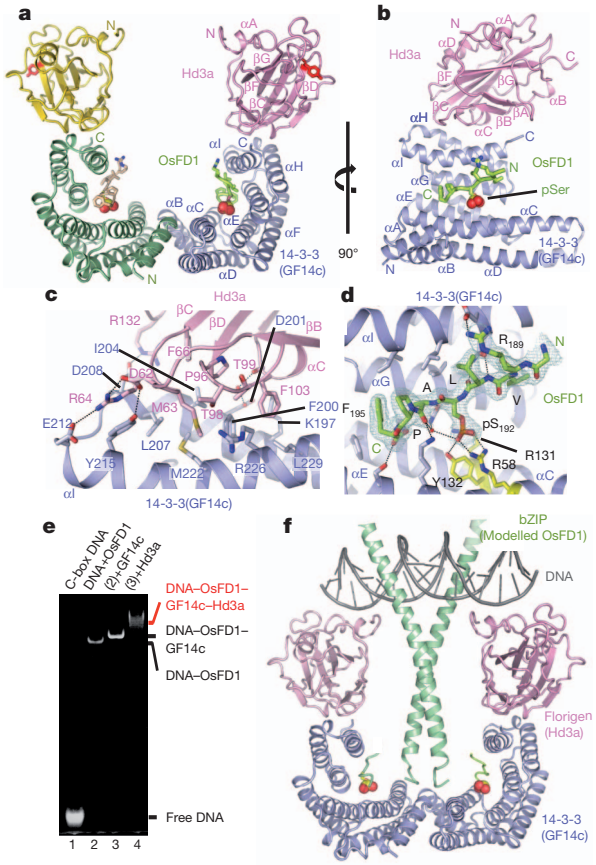


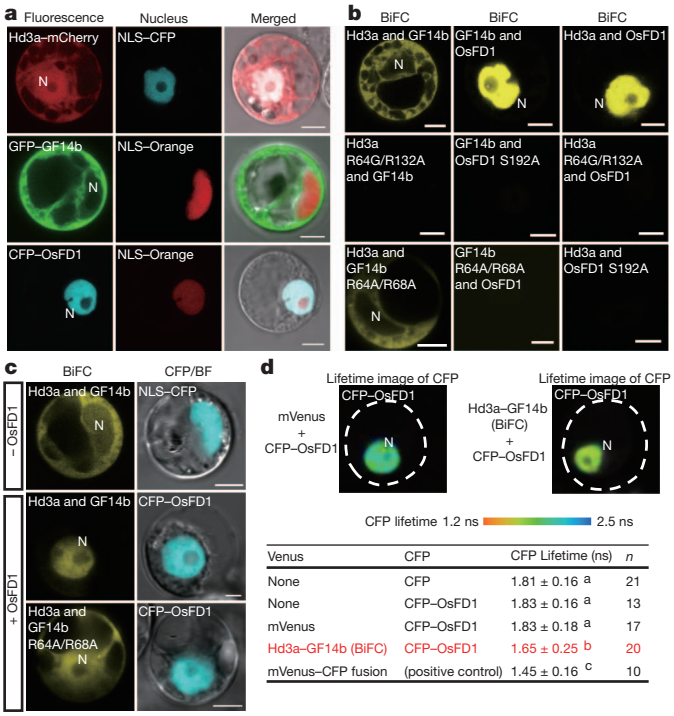

**Figure 2 | Structure of the florigen activation complex.** **a**, The FAC hetero-hexameric structure, in a ribbon representation, composed of two Hd3a molecules (gold and magenta), a GF14c dimer (dark green and blue) and two OsFD1 peptides (khaki and light green). Hd3a Y87 residues are shown by red sticks. **b**, The same structure as **a**, rotated 90° around the vertical axis; the right half of the complex depicted in **a** is shown. pSer, phosphoserine. **c**, Expanded view of the interface between Hd3a (magenta) and GF14c (blue). **d**, Expanded view of the interface between GF14c (blue) and OsFD1 (green). **e**, Interaction between FAC components and DNA, examined by gel-shift assays. **f**, Modelled structure of the FAC–promoter-DNA complex. Hd3a, GF14, OsFD1 and DNA are coloured in magenta, blue, green and grey, respectively.

The ability of OsFD1 and the FAC to bind to DNA was examined by gel-shift assays. A DNA oligomer of 22 base pairs containing the C-box (GACGTC) element from the *AP1* promoter was used for this experiment; the element also occurs in the promoter of *OsMADS15*, which is activated by Hd3a (Supplementary Fig. 16). We observed supershifts of DNA upon addition of OsFD1, GF14c and Hd3a to the *AP1* promoter oligomer, but there was no supershift when GF14c was omitted (Fig. 2e and Supplementary Fig. 17). These data indicate that the FAC forms a stable complex with the promoter regions of *AP1* and its homologues. As shown in Fig. 2f, a model of the FAC–DNA complex was built (Supplementary Text 4). In this model, OsFD1 does not contact Hd3a, but the two Hd3a molecules in this trident-shaped structure may contact DNA. Experimentally, OsFD1 and the FAC bind directly to DNA (Fig. 2e). Two Hd3a molecules therefore seem to become anchored to the promoter DNA through interactions involving 14-3-3 and bZIP proteins, and this anchored Hd3a may serve as the specific trigger of the flowering signal (Supplementary Text 4)[16].

We then examined the subcellular localization and interaction of these three proteins in rice protoplasts (Supplementary Text 5). We detected fluorescence from Hd3a tagged with mCherry in both the cytoplasm and the nucleus; from GFP-tagged GF14b in the cytoplasm and also weakly in the nucleus; and from OsFD1 tagged with mCerulean (CFP), in the nucleus (Fig. 3a and Supplementary Table 2). The Hd3a–GF14b interaction was monitored by bimolecular fluorescence complementation (BiFC) assays and was detected in the cytoplasm, where GF14b was mainly localized (Fig. 3b and Supplementary Figs 20 and 21). Notably, the GF14b–OsFD1 BiFC signal was mainly detected in the nucleus, whereas GF14b alone was always detected predominantly in the cytoplasm. BiFC assays also detected an association between Hd3a and OsFD1 in the nucleus, as reported for the FT–FD interaction[6,7]. However, the Hd3a–OsFD1 interaction was abolished in the Hd3a R64G/R132A and OsFD1 S192A mutants, both of which also lost the ability to interact with GF14b. To assess the influence of OsFD1 on the subcellular localization of Hd3a, GF14b and the Hd3a–GF14b complex, we combined the above assays with CFP–OsFD1 co-expression. When CFP–OsFD1 was co-expressed, GFP–GF14b, Hd3a–mCherry and the Hd3a–GF14b complex became clearly concentrated in the nucleus (Fig. 3c, Supplementary Figs 22 and 23 and Supplementary Table 2). Taken together, these results indicate that the Hd3a–GF14b complex is translocated from the cytoplasm to the nucleus, where it forms a larger protein complex with OsFD1. To obtain further evidence for the existence of the ternary complex in living cells, we conducted BiFC-based fluorescence resonance energy transfer (FRET)–fluorescence lifetime imaging microscopy (FLIM) analysis[17,18]. We found that the lifetime of CFP fluorescence in cells co-transformed with CFP–OsFD1 and Hd3a–GF14b BiFC vectors was significantly shorter than that in control cells (Fig. 3d and

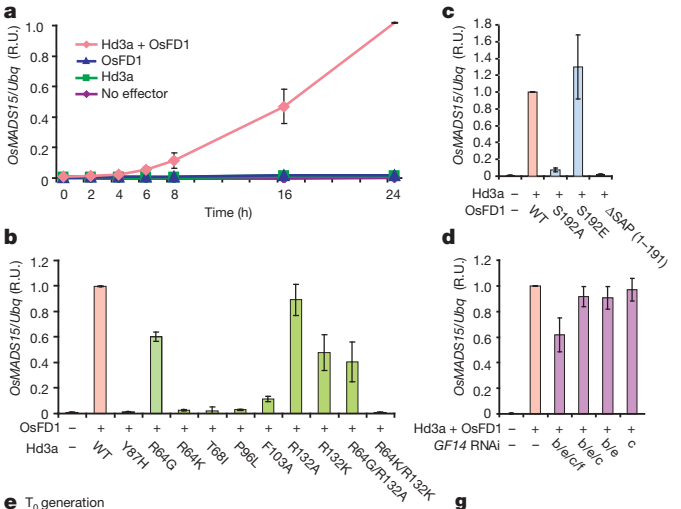


©2011 Macmillan Publishers Limited. All rights reserved




| Venus | CFP | CFP Lifetime (ns) | *n* |
|---|---|---|---|
| None | CFP | 1.81 ± 0.16 [a] | 21 |
| None | CFP–OsFD1 | 1.83 ± 0.16 [a] | 13 |
| mVenus | CFP–OsFD1 | 1.83 ± 0.18 [a] | 17 |
| Hd3a–GF14b (BiFC) | CFP–OsFD1 | 1.65 ± 0.25 [b] | 20 |
| mVenus–CFP fusion | (positive control) | 1.45 ± 0.16 [c] | 10 |

**Figure 3 | The Hd3a–GF14b complex enters the nucleus to interact with OsFD1.** **a**, Confocal images of cells expressing Hd3a–mCherry, GFP–GF14b and CFP–OsFD1. Nuclear marker proteins (NLS–CFP or NLS–Orange) were co-expressed. **b**, BiFC assays showing interactions of Hd3a–GF14b (left), GF14b–OsFD1 (centre) and Hd3a–OsFD1 (right). Venus fluorescence in cells expressing the indicated proteins tagged with the N- or C-terminal halves of Venus (see also Supplementary Fig. 16) is shown. **c**, BiFC assay for Hd3a–GF14b interaction and OsFD1 co-expression. CFP/BF, merged image of CFP fluorescence and bright field. **d**, FLIM measurements of Hd3a–GF14b–OsFD1 interaction. CFP fluorescence lifetime images of representative cells are shown, with fluorescence lifetimes tabulated below. In the table, CFP lifetime values followed by 'b' or 'c' are significantly different ($P < 0.05$). Scale bars, 5 μm.




**e** $T_0$ generation

| Genotype | Days to flowering | *n* | *P* value |
|---|---|---|---|
| WT | 55.7 ± 9.4 | 9 | |
| *prolC:Hd3a–GFP* | 19.6 ± 12.3 ** | 13 | $2.0 \times 10^{-7}$ |
| *prolC:Hd3a R64G–GFP* | 44.2 ± 7.8 * | 6 | 0.02 |
| *prolC:Hd3a R64G/R132A–GFP* | 58.8 ± 9.3 | 21 | 0.42 |
| *prolC:Hd3a R64K/R132K–GFP* | 61.9 ± 9.7 | 15 | 0.13 |
| *prolC:Hd3a F103A–GFP* | 54.7 ± 7.9 | 7 | 0.83 |
| *pUbq:OsFD1* | 49.7 ± 9.6 | 7 | 0.23 |
| *pUbq:OsFD1 S192A* | 61.7 ± 2.9 | 3 | 0.12 |
| *pUbq:OsFD1 S192E* | 36.1 ± 4.8 ** | 14 | 0.00013 |

**f** $T_1$ generation

| Genotype | Days to flowering | *n* | *P* value |
|---|---|---|---|
| Experiment 1 | | | |
| WT | 63.1 ± 2.9 | 12 | |
| *OsFD1* RNAi | 66.2 ± 2.7 * | 10 | 0.016 |
| *pUbq:OsFD1 S192A* | 61.3 ± 2.1 | 7 | 0.14 |
| *Hd3a RNAi* | 85.2 ± 2.0 ** | 9 | $1.9 \times 10^{-14}$ |
| *pRPP16:Hd3a–GFP* | 51.5 ± 3.8 ** | 4 | 0.004 |
| Experiment 2 | | | |
| WT | 59.3 ± 2.0 | 6 | |
| *prolC:Hd3a–GFP*[*1] | 23.9 ±11.1 ** | 7 | 0.00011 |
| *prolC:Hd3a R64G–GFP* | 59.8 ± 3.8 | 8 | 0.79 |
| *prolC:Hd3a R64G/R132A–GFP* | 59.6 ± 1.5 | 5 | 0.81 |

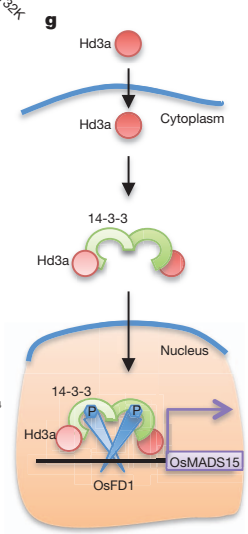



**Figure 4 | 14-3-3 interaction is required for *OsMADS15* activation by Hd3a and OsFD1.** **a**, Induction of *OsMADS15* requires both Hd3a and OsFD1 in suspension-cultured cells. R.U., relative units. **b**, Effects of Hd3a mutations on *OsMADS15* activation. **c**, Effects of OsFD1 mutations on *OsMADS15* activation. **d**, Effects of simultaneous knockdown of *GF14b*, *GF14c*, *GF14e* and *GF14f* on *OsMADS15* activation. Data represent three experiments; all error bars represent mean ± s.e.m. **e**, **f**, Flowering times of transgenic plants of $T_0$ (**e**) and $T_1$ (**f**) generations under short-day conditions. *, $P < 0.05$; **, $P < 0.01$; [1], $T_0$ generation was examined for *prolC:Hda3–GFP*. **g**, Model of FAC formation. Once Hd3a florigen enters a shoot apical cell, it initially binds 14-3-3 proteins in the cytoplasm. When the Hd3a–14-3-3 complex enters the nucleus, it forms a complex with OsFD1, which is retained in the nucleus and activates *OsMADS15* transcription, leading to floral induction.

Supplementary Fig. 24). These results indicate that the Hd3a–GF14b–OsFD1 tripartite protein complex is indeed formed in rice cells (Supplementary Text 5).

*AP1* transcription is dependent on FT and FD[6,7] and, in rice, RNA interference (RNAi)-mediated knockdown of *Hd3a* strongly attenuates the expression of *OsMADS15* (ref. 9). Therefore, we next examined how the interactions among Hd3a, 14-3-3 and OsFD1 affect *OsMADS15*, using rice protoplasts (Supplementary Text 6 and 7). The level of *OsMADS15* messenger RNA increased with time after *Hd3a* and *OsFD1* co-transformation (Fig. 4a), whereas neither *Hd3a* nor *OsFD1* alone induced *OsMADS15* expression. To evaluate the importance of the 14-3-3 interaction, various Hd3a mutants which are defective in 14-3-3 binding (Supplementary Fig. 8) were co-expressed with OsFD1 (Fig. 4b). The results indicate that the interaction of Hd3a with 14-3-3 is critical for *OsMADS15* expression. Mutations of the SAP motif in OsFD1 (ΔSAP and S192A), but not a phospho-mimicking substitution (S192E), greatly reduced *OsMADS15* transcription (Fig. 4c). Simultaneous knockdown by RNAi of four *GF14* genes also reduced *OsMADS15* mRNA levels (Fig. 4d and Supplementary Fig. 25). Our results support a model in which a protein complex consisting of Hd3a, 14-3-3 and OsFD1 induces *OsMADS15* expression, with the phosphorylated serine residue S192 in the OsFD1 SAP motif being recognized by 14-3-3 to facilitate the association between OsFD1 and Hd3a.

Hd3a–GFP protein expressed from the phloem-specific *rolC* promoter in transgenic plants moves to the shoot apex and strongly accelerates flowering[5]. Compared with plants expressing *prolC:Hd3a–GFP*, plants expressing all the Hd3a mutants with reduced affinity for 14-3-3 lost the accelerated flowering (Fig. 4e, f, Supplementary Fig. 26 and Supplementary Tables 3 and 4). These results indicate that the binding of 14-3-3 to Hd3a is important for flowering. *OsFD1* RNAi plants flowered later than wild-type plants, reminiscent of the *Arabidopsis fd* and maize *dlf1* mutant phenotypes[6,7,19], and plants expressing *OsFD1 S192E* under the ubiquitin (*Ubq*) promoter had accelerated flowering, indicating that OsFD1 phosphorylation is a rate-limiting step for FAC formation, and hence for flowering (Fig. 4e, f, Supplementary Figs 27 and 28 and Supplementary Tables 3 and 4). None of the transgenic plants with modified 14-3-3 expression had altered flowering times (Supplementary Figs 27–30 and Supplementary Table 5), probably because of high redundancy and high expression of *14-3-3* genes[11].

In our model (Fig. 4g), 14-3-3 proteins act as intracellular receptors for Hd3a (florigen), which is transported from leaves to the shoot apex. Once Hd3a enters cells in the shoot apex, it initially binds 14-3-3 proteins in the cytoplasm. At this developmental stage, OsFD1 is being



©2011 Macmillan Publishers Limited. All rights reserved

expressed in the shoot apex cells, and when the Hd3a–14-3-3 complex enters the nucleus, it forms a ternary complex with OsFD1 that is retained in the nucleus. Alternatively, newly translated OsFD1 in the cytoplasm may interact there with Hd3a–14-3-3 to form the FAC, which is then translocated into the nucleus. The FAC then activates transcription of *OsMADS15*, leading to floral induction. One notable feature of the proposed mechanism is that incoming florigen can be efficiently trapped in the cytoplasm of shoot apex cells by 14-3-3 proteins, which are highly expressed in these cells[11], and transported to the nucleus for transcriptional activation of *OsMADS15*. Our model provides a mechanistic basis for the proposed participation of florigen in processes other than flowering[20–22]. Therefore, our findings offer clues for improving various agriculturally important traits by manipulating florigen and other interacting regulators.

## METHODS SUMMARY

Screening of Hd3a interactors using a leaf cDNA library was performed as described previously[11]. For recombinant protein expression, the pCold-GST[23] or pGEX 6P-3 vectors (GE Healthcare) were used. All NMR spectra were measured using AV500, DRX800 and AVANCE 800 spectrometers (Bruker) at 303 K, and were processed using NMRpipe software[24]. Hd3a (K31A/E57A) and Hd3a (wild-type), labelled with $^{13}C$ and $^{15}N$, were used for NMR. Isothermal titration calorimetry (ITC) experiments were performed at 298 K on a VP-ITC Micro Calorimeter (Microcal). For immunoprecipitation, samples of the shoot apical region were dissected under a microscope. A μMACS GFP isolation kit (Miltenyi Biotec) was used to capture GFP in extracts and the eluted proteins were then used for western blotting. Hd3a (residues 6–170) C43L/C109S/C166S and GF14c (residues 1–235) were used for crystallization. Hd3a–GF14c–OsFD1 crystals were produced by soaking the Hd3a–GF14c crystals with OsFD1 peptide. Crystallographic data were collected on the BL-5A beamline at the Photon Factory and on the BL41XU at SPring-8. The crystal structure was solved and refined using appropriate software[25–27]. The FAC–DNA complex was modelled by superimposing the structure of the mouse CREB bZIP–C-box DNA complex (Protein Data Bank accession code 1DH3) onto that of the FAC[28]. Rice cultured cell protoplasts were transformed as previously described[29]. Confocal laser scanning microscopy observations and BiFC assays using protoplasts were conducted as previously described[29]. In the FLIM experiments, the CFP lifetime in a pixel was calculated from the fluorescence decay curve produced for every pixel, and the frequency distribution of lifetimes in a cell was then calculated. Generation of transgenic rice plants and maintenance of cultured cells were conducted as previously described[5]. Rice (*Oryza sativa* L. subspecies Japonica) variety Norin 8 was used as wild type. Extraction of RNA and reverse transcriptase PCR (RT–PCR) were performed as described[9]. For stable and transient RNAi experiments, pANDA and pANDA-mini vectors were used[30].

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 15 February; accepted 6 June 2011.**
**Published online 31 July 2011.**

1. Chailakhyan, M. H. On the hormonal theory of plant development. *C. R. Acad. Sci. U.R.S.S.* **12,** 443–447 (1936).
2. Zeevaart, J. A. Leaf-produced floral signals. *Curr. Opin. Plant Biol.* **11,** 541–547 (2008).
3. Kobayashi, Y. & Weigel, D. Move on up, it's time for change—mobile signals controlling photoperiod-dependent flowering. *Genes Dev.* **21,** 2371–2384 (2007).
4. Turck, F., Fornara, F. & Coupland, G. Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. *Annu. Rev. Plant Biol.* **59,** 573–594 (2008).
5. Tamaki, S., Matsuo, S., Wong, H. L., Yokoi, S. & Shimamoto, K. Hd3a protein is a mobile flowering signal in rice. *Science* **316,** 1033–1036 (2007).
6. Abe, M. *et al.* FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science* **309,** 1052–1056 (2005).
7. Wigge, P. A. *et al.* Integration of spatial and temporal information during floral induction in *Arabidopsis*. *Science* **309,** 1056–1059 (2005).
8. Komiya, R., Yokoi, S. & Shimamoto, K. A gene network for long-day flowering activates RFT1 encoding a mobile flowering signal in rice. *Development* **136,** 3443–3450 (2009).
9. Komiya, R., Ikegami, A., Tamaki, S., Yokoi, S. & Shimamoto, K. Hd3a and RFT1 are essential for flowering in rice. *Development* **135,** 767–774 (2008).
10. Kaufmann, K. *et al.* Orchestration of floral initiation by APETALA1. *Science* **328,** 85–89 (2010).
11. Purwestri, Y. A., Ogaki, Y., Tamaki, S., Tsuji, H. & Shimamoto, K. The 14-3-3 protein GF14c acts as a negative regulator of flowering in rice by interacting with the florigen Hd3a. *Plant Cell Physiol.* **50,** 429–438 (2009).
12. Pnueli, L. *et al.* Tomato SP-interacting proteins define a conserved signaling system that regulates shoot architecture and flowering. *Plant Cell* **13,** 2687–2702 (2001).
13. Lifschitz, E. *et al.* The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. *Proc. Natl Acad. Sci. USA* **103,** 6398–6403 (2006).
14. Yaffe, M. B. *et al.* The structural basis for 14-3-3:phosphopeptide binding specificity. *Cell* **91,** 961–971 (1997).
15. Wilker, E. W., Grant, R. A., Artim, S. C. & Yaffe, M. B. A structural basis for 14-3-3σ functional specificity. *J. Biol. Chem.* **280,** 18891–18898 (2005).
16. Ahn, J. H. *et al.* A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. *EMBO J.* **25,** 605–614 (2006).
17. Kwaaitaal, M., Keinath, N. F., Pajonk, S., Biskup, C. & Panstruga, R. Combined bimolecular fluorescence complementation and Förster resonance energy transfer reveals ternary SNARE complex formation in living plant cells. *Plant Physiol.* **152,** 1135–1147 (2010).
18. Shyu, Y. J., Suarez, C. D. & Hu, C. D. Visualization of AP-1–NF-κβ ternary complexes in living cells by using a BiFC-based FRET. *Proc. Natl Acad. Sci. USA* **105,** 151–156 (2008).
19. Muszynski, M. G. *et al. delayed flowering1* encodes a basic leucine zipper protein that mediates floral inductive signals at the shoot apex in maize. *Plant Physiol.* **142,** 1523–1536 (2006).
20. Bohlenius, H. *et al.* CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* **312,** 1040–1043 (2006).
21. Shalit, A. *et al.* The flowering hormone florigen functions as a general systemic regulator of growth and termination. *Proc. Natl Acad. Sci. USA* **106,** 8392–8397 (2009).
22. Krieger, U., Lippman, Z. B. & Zamir, D. The flowering gene *SINGLE FLOWER TRUSS* drives heterosis for yield in tomato. *Nature Genet.* **42,** 459–463 (2010).
23. Hayashi, K. & Kojima, C. pCold-GST vector: a novel cold-shock vector containing GST tag for soluble protein production. *Protein Expr. Purif.* **62,** 120–127 (2008).
24. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6,** 277–293 (1995).
25. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D* **61,** 458–464 (2005).
26. Vagin, A. A. *et al.* REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D* **60,** 2184–2195 (2004).
27. Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
28. Schumacher, M. A., Goodman, R. H. & Brennan, R. G. The structure of a CREB bZIP·Somatostatin CRE complex reveals the basis for selective dimerization and divalent cation-enhanced DNA binding. *J. Biol. Chem.* **275,** 35242–35247 (2000).
29. Kawano, Y. *et al.* Activation of a Rac GTPase by the NLR family disease resistance protein Pit plays a critical role in rice innate immunity. *Cell Host Microbe* **7,** 362–375 (2010).
30. Miki, D. & Shimamoto, K. Simple RNAi vectors for stable and transient suppression of gene function in rice. *Plant Cell Physiol.* **45,** 490–495 (2004).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank Y. Takahashi for the anti-14-3-3 antibody and S. Takayama for the BiFC vectors. We are grateful to I. Smith for critical reading of the manuscript. We also thank N. Inada for FRET measurement; Y. Nishio, Y. Ohno, E. Kawano and M. Kanda for technical assistance; Y. Tamaki, Y. Konomi and J. Naritomi for rice transformation; M. Yoneyama, H. Kinoshita and A. Yasuba for sample preparation; J. Tsukamoto for mass spectroscopy analysis; H. Fukada for ITC measurement; K. Takeshita and Y. Hara for crystallization; and members of the Laboratory of Plant Molecular Genetics and the Laboratory of Biophysics at Nara Institute of Science and Technology (NAIST) for discussions. We thank beamline staff of BL41XU and BL44XU at SPring-8 and Photon Factory for data collection. We also thank Academia Sinica and National Synchrotron Radiation Research Center (Taiwan) for use of the MX225-HE detector at BL44XU. This work was supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) through the Target Proteins Research Program, by the Global COE Program at NAIST, by Grants-in-Aid for Scientific Research to I.O., C.K. and K.S., by Grants-in-Aid for Scientific Research on Priority Areas to K.S. and by the Program for Promotion of Basic and Applied Researches for Innovations in Bio-oriented Industry from Bio-oriented Research Advancement Institution (BRAIN) to H.T.

**Author Contributions** K.T., I.O., H.T., K.F., K.H., T.Y., M.Y., C.N., Y.A.P., S.T., Y.O. and C.S. performed experiments; I.O. conceived the idea that 14-3-3 mediates the interaction between Hd3a and OsFD1; K.T., I.O., H.T., A.N., C.K. and K.S. analysed data; and K.T., I.O., H.T., C.K. and K.S. wrote the paper. C.K. and K.S. contributed equally to this work.

**Author Information** The atomic coordinates, structure factors and NMR data have been deposited in the Protein Data Bank (accession number 3AXY) and Biological Magnetic Resonance Bank (accession number 11442). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.K. (kojima@protein.osaka-u.ac.jp) or K.S. (simamoto@bs.naist.jp).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Plant materials and growth conditions.** Rice (*Oryza sativa* L. subspecies Japonica) variety Norin 8 was used as wild type. *pHd3a:Hd3a–GFP*, *35S:GFP* and *prolC:Hd3a–GFP* transgenic rice plants were described previously[5,31]. Transgenic rice plants were generated using *Agrobacterium*-mediated transformation of rice calli, as previously described[32], and hygromycin-resistant plants were regenerated from the transformed callus. Transgene integration was further confirmed by PCR amplification of the hygromycin phosphotransferase gene in genomic DNA extracted from regenerated plants. Plants were grown in climate chambers at 80% humidity, under short-day conditions with daily cycles of 10 h of light at 30 °C and 14 h of dark at 25 °C. Light was provided by fluorescent white-light tubes (400–700 nm, 100 μmol $m^{-2}$ $s^{-1}$). Flowering time was measured as the number of days to the heading stage after $T_0$ transgenic plants were transferred to short-day conditions. Rice suspension-cultured cells were maintained as described previously[33].

**Yeast two-hybrid assays.** Screening of Hd3a interactors was performed as described previously[11]. Briefly, yeast two-hybrid libraries were prepared using total RNA extracted from wild-type leaf blades which were harvested 35–40 days after sowing. After cDNA synthesis using a cDNA synthesis kit (Stratagene), the cDNAs were inserted into the pVP16 vector[34] and introduced into the yeast L40 strain. As a bait, the full-length Hd3a open reading frame was cloned into the pBTM116 vector[35]. Screening was performed on SC medium lacking histidine and containing 2.5 mM 3-aminotriazole (3-AT).

For interaction assays, pBTM116 and pVP16 were converted to the Gateway destination vectors pBTM116-GW and pVP16-GW, using the Gateway vector conversion system (Invitrogen) according to the manufacturer's instructions. Yeast cells were grown at 30 °C for 5 days on SC medium without uracil, tryptophan, leucine and histidine, containing added histidine or 1–10 mM 3-AT. The concentration of 3-AT was determined by the bait–prey combination.

**Co-immunoprecipitation assay.** Samples of the shoot apical region were collected by microscopic dissection and total proteins were extracted. These extracts were used for immunoprecipitation using the μMACS GFP isolation kit (Miltenyi Biotec). After calibrating the amount of eluted protein, the eluates were separated by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and subjected to immunoblotting using anti-GFP antibody (Abcam) and anti-14-3-3 antibody (provided by Y. Takahashi). After washing with TBST buffer, the membranes were incubated for 1 h with anti-rabbit IgG conjugated to horseradish peroxidase (GE Healthcare). Detection was performed using enhanced chemiluminescence (ECL) detection reagents (GE Healthcare).

**Protoplast transformation.** Transformation of rice Oc protoplasts was performed as described previously[33]. For transient expression analysis, 8 μg of *Hd3a* expression vectors and 16 μg of *OsFD1* expression vectors were introduced into 500 μl of protoplast suspension at a concentration of $2 \times 10^7$ protoplasts $ml^{-1}$ by the PEG-mediated transformation method. After a 24-h incubation at 30 °C, the protoplast suspension was centrifuged and the cell pellet was frozen at −80 °C for RNA extraction. For transient RNAi analysis, 24 μg of *GF14* RNAi vector was introduced and the transformed protoplasts were incubated for 72 h to reduce the amount of carried-over GF14 proteins.

**Subcellular localization and bimolecular fluorescence complementation.** The Hd3a, GF14b, OsFD1 and β-glucuronidase coding regions were cloned into fluorescent protein expression vectors or BiFC vectors and purified using the Purelink Plasmid Midiprep kit (Invitrogen). The amount of plasmids used for transformation was as follows. For the GFP–GF14b and CFP–OsFD1 co-expression experiment, we co-transformed 1 μg of GFP–GF14b or Hd3a–mCherry expression plasmid and 10 μg of NLS–CFP, CFP–OsFD1 or CFP–OsFD1 S192A expression plasmid. For BiFC experiments, 5 μg of each Venus-N-terminal- and Venus-C-terminal-tagged protein expression vector was co-transformed, except for the GF14b–OsFD1 interaction assay, in which the quantities were 10 μg Vc–GF14b or Vc–GF14b R64A/R68A and 1 μg Vn–OsFD1. When CFP–OsFD1 was co-expressed in the BiFC experiment, 10 μg of NLS–CFP or CFP–OsFD1 expression plasmid was used. The mCherry expression plasmid was introduced simultaneously as a marker for transformation efficiency. To quantify protein–protein interactions from the BiFC assay, the fluorescence intensities of Venus (BiFC) and mCherry (transformation marker) from about 20 cells were measured under the same microscope settings, and the Venus/mCherry value was calculated[29]. In our experimental settings, we could recognize reliable BiFC signal in cells showing Venus/mCherry ratios of >0.33 for the Hd3a–GF14b interaction and >0.4 for the GF14b–OsFD1 and Hd3a–OsFD1 interactions. Therefore, the number of cells showing ratios exceeding these values was scored.

We examined the degree of nuclear accumulation of the BiFC signal arising from the Hd3a–GF14b interaction in cells co-expressing OsFD1. To do this, we established a method to calculate nuclear accumulation. First, we measured the fluorescence intensities of Venus and mCherry in the nuclei of transformed cells. Next, we calculated the value of (Venus in nucleus/mCherry in nucleus)/(Venus in whole cell/mCherry in whole cell), which indicates the ratio of nuclear accumulation of Venus. Finally, we compared these values and the corresponding confocal images from each cell, and found that values >1.2 could be recognized as nuclear localization, values of 0.8–1.2 (experiment 1) or 1.0–1.2 (experiment 2), as localization in both nucleus and cytosol, and values <0.8 (experiment 1) or <1.0 (experiment 2), as cytosolic localization, compared with mCherry distribution.

**FLIM for FRET detection.** A Leica TCS SP5 confocal laser scanning microscope and a time-correlated single-photon counting module (SPC-830; Becker & Hickl) were used for FLIM. A diode pulsed laser was used to excite CFP and the emission from 420 to 490 nm was collected in 64 × 64 pixel mode. The CFP lifetime in a pixel was calculated from the fluorescence decay curve produced for every pixel, and the frequency distribution of lifetime was then calculated by combining all pixel data from single-cell observations. The mode of frequency distribution was defined as the lifetime of CFP in the cell. Because CFP–OsFD1 was localized in the nucleus, the lifetime in the nucleus was measured for all experiments.

**RNA extraction and real-time PCR analysis.** Leaf blades of wild-type and transgenic rice plants were harvested at Zeitgeber time (ZT) 0-1. Total RNA was extracted using TRizol reagent (Invitrogen) according to the manufacturer's protocol. cDNA was synthesized from 1 μg or 2 μg of RNA, using a 21-nucleotide oligo dT primer and Superscript II reverse transcriptase (Invitrogen). cDNA (1 μl) was used for quantitative analysis of gene expression using SYBR Green PCR master mix (Applied Biosystems) with gene-specific primers. Data were collected using the ABI PRISM 7000 sequence detection system in accordance with the manufacturer's instruction manual. The primers used for *ubiquitin* and *OsMADS15* were described previously[9].

**Plasmid construction.** Full-length cDNAs for *GF14a* (Os08g0480800), *GF14b* (Os04g0462500), *GF14c* (Os08g0430500), *GF14d* (Os11g0546900), *GF14e* (Os02g0580300), *GF14f* (Os03g0710800), *GF14g* (Os01g0209200), *GF14h* (Os11g0609600), *OsFD1* (Os09g0540800), *OsKANADI1* (Os02g0696900) and *OsBIP116b* (Os07g0185432) were cloned by RT–PCR. The forward and reverse primers were designed on the basis of gene information in the rice annotation DNA database (http://rapdb.dna.affrc.go.jp/). The coding regions were PCR-amplified and ligated into the pENTR/D-TOPO cloning vector (Invitrogen) to obtain entry clones. Primer sequences are available on request.

Introduction of amino-acid substitutions or deletions into *GF14*, *OsFD1*, *Hd3a*, *OsKANADI1* and *OsBIP116b* was performed by PCR with KOD FX DNA polymerase (Toyobo). The PCR-amplified fragments were cloned into pENTR/D-TOPO, as above. Primer sequences are available on request.

For RNAi vectors, parts of the 3′ untranslated regions of *GF14b* RNA (237 nucleotides), *GF14c* RNA (310 nucleotides), *GF14e* RNA (242 nucleotides) and *GF14f* RNA (303 nucleotides) were reverse-transcribed, PCR-amplified and concatenated by PCR. The PCR-amplified fragments were cloned into pENTR/D-TOPO, as above.

The parental vectors for transient expression assays, pGII pUbq GW-T7 and pGII pUbq HA-GW, were constructed by inserting the *ubiquitin* promoter from maize, the *NOS* termination region[30] and the *attR* recombination region with T7 or HA tag regions[36] into pGreen II vector[37]. The *pUbq:Hd3a–mCherry* transient expression vector and the *pHd3a:Hd3a–GFP* and *prolC:Hd3a–GFP* binary vectors[5] were described previously. The *prolC:Hd3a R64G–GFP* and *prolC:Hd3a R64G/R132A–GFP* binary vectors were generated as previously described[5], using mutant *Hd3a* cDNA. To generate other vectors, the *Hd3a*, *GF14*, *OsFD1*, *OsKANADI1* and *OsBIP116b* mutant genes in pENTR/D-TOPO vectors were transferred into various vectors by the Gateway recombination method with Gateway LR clonase II (Invitrogen), depending on the experiment. These included pBTM116-GW and pVP16-GW for the yeast two-hybrid assay, pGII pUbq GW-T7, pGII pUbq HA-GW, p35S-GFP-GW, p35S-mCerulean-GW, p35S-Vn-GW, p35S-Vc-GW, p35S-GW-Vn, p35S-GW-Vc, pUbq-GW and the pANDA mini vector[30] for transient expression assays, p2K-GW binary vector (for transgenic plants expressing genes under the maize *ubiquitin* promoter) and the pANDA RNAi vector[30] for transgenic plants. The SV40 nuclear localization sequence was fused with *GF14b* and *GF14e* by PCR. Details of these plasmid constructions are available on request.

**Protein expression and purification.** The Hd3a gene was cloned into the pCold-GST vector[23] and expressed as a GST fusion protein in *Escherichia coli* BL21 Rosetta (DE3) (Novagen). Cells were grown in minimal medium containing 0.5 $g\,l^{-1}$ $^{15}N$-ammonium chloride for NMR titration and GST pull-down assay, or 0.5 $g\,l^{-1}$ $^{15}N$-ammonium chloride and 2 $g\,l^{-1}$ $^{13}C$-glucose for NMR assignment. The expressed GST-fused protein was purified by glutathione sepharose 4B resin (GE Healthcare). After digestion of the GST tag using GST-HRV 3C protease, Hd3a was purified by gel filtration chromatography using a Superdex75 column (GE Healthcare). To prepare Hd3a mutants for NMR measurement and GST pull-down assay, expression plasmids were constructed using a Quikchange site-directed mutagenesis kit (Stratagene) and the expressed proteins were purified by essentially the same method as described above. The purified Hd3a protein was dialysed against

©2011 Macmillan Publishers Limited. All rights reserved

50 mM potassium phosphate buffer (pH 6.8) containing 50 mM KCl, 1 mM dithiothreitol (DTT), 5 mM EDTA and 5 mM EGTA for 30 min at 4 °C, and then against 50 mM potassium phosphate buffer (pH 6.8) containing 50 mM KCl and 1 mM DTT overnight at 4 °C. For crystallization, Hd3a (6–170) C43L/C109S/C166S was cloned into the pCold-GST vector, expressed as a GST fusion protein in *E. coli* BL21 Rosetta (DE3) growing in LB medium, and purified as described above.

The gene encoding OsGF14c was cloned into the pGEX 6P-3 vector (GE Healthcare). Full-length GF14 and C-terminally truncated GF14 (1–235) were expressed as GST-fused proteins in *E. coli* BL21 Rosetta (DE3) and *E. coli* BL21 (DE3), respectively. To prepare expression plasmids of GF14 mutants, the Quikchange site-directed mutagenesis kit was used. For the NMR titration experiment and crystallization, GF14 proteins were purified by glutathione sepharose 4B resin and a Superdex 200 column, followed by digestion of the GST tag.

The *OsFD1* gene was cloned into the pCold-GST vector. N-terminally truncated OsFD1 (147–195) S192E and full-length OsFD1 S192E were expressed as GST fusion proteins in *E. coli* BL21 Rosetta (DE3) for GST pull-down, NMR, circular dichroism (CD) and isothermal titration calorimetry (ITC) measurements. For the DNA binding assay, OsFD1 (124–195) S192E was expressed as a GST fusion protein. Proteins were purified by the same method used for the Hd3a proteins.

**GST pull-down assay.** Purified GST-tagged proteins (10 nmol) were adsorbed onto 10 µl of glutathione sepharose 4B resin. For the negative control, 10 nmol of purified GST protein was used in place of the GST-tagged protein. After washout of excess GST-tagged proteins or GST, 1 nmol of purified proteins lacking GST tags was added. After incubation at room temperature for 15 min, unbound proteins were washed away. Bound proteins were eluted with GST-tagged proteins using 20 µl of phosphate buffer (pH 6.8) containing 50 mM KCl, 50 mM glutathione and 1 mM DTT. Eluted fractions were separated by SDS–PAGE and visualized by silver staining. The amount of bound protein was quantified by ImageJ software (NIH).

**NMR experiments.** All NMR spectra were measured using AV500, DRX800 and AVANCE 800 spectrometers (Bruker) at 303 K and processed using NMRpipe software[24]. Non-uniformly sampled spectra were processed by Rowland NMR Toolkit[38] software. All NMR spectra were analysed using Sparky software[39].

Backbone chemical shifts of Hd3a (K31A/E57A) were assigned in a conventional manner. The $^{13}C/^{15}N$-labelled sample was prepared at a concentration of 0.2 mM in 50 mM potassium phosphate buffer (pH 6.8) containing 100 mM KCl and 1 mM DTT (93% $H_2O$, 7% $D_2O$ (v/v)), and non-uniformly sampled HNCA, HN(CO)CA, HN(CA)CB and HN(COCA)CB spectra[40] were recorded. Backbone amide chemical shifts of Hd3a (wild-type) were then assigned by comparing the $^1H$-$^{15}N$ HSQC spectrum[41] with that of Hd3a (K31A/E57A), and these assignments were confirmed by checking sequence connectivity in the non-uniformly sampled HNCA and HN(CO)CA spectra.

$^1H$-$^{15}N$ HSQC titration experiments of 0.1 mM or 0.15 mM $^{15}N$-labelled Hd3a with unlabelled GF14 were performed in 20 mM MES buffer (pH 6.8) containing 50 mM KCl and 1 mM DTT. GF14 was added at Hd3a:GF14 molar ratios of 1:0.1, 1:0.5 and 1:1. $^1H$-$^{15}N$ HSQC titration experiments of $^{15}N$-labelled Hd3a with an unlabelled S192E mutant mimicking phosphorylated OsFD1 were performed in 50 mM potassium phosphate buffer (pH 6.8) containing 100 mM KCl and 1 mM DTT. OsFD1 was added at Hd3a:GF14 molar ratios of 1:1 and 1:2. Interactions between Hd3a mutants and GF14 were monitored by comparing $^1H$-$^{15}N$ HSQC spectra of ~0.1 mM $^{15}N$-labelled Hd3a in the presence and absence of unlabelled GF14 at an Hd3a:GF14 molar ratio of 1:0.5, with 50 mM potassium phosphate buffer (pH 6.8) containing 50 mM KCl and 1 mM DTT.

$^1H$-$^{15}N$ heteronuclear nuclear Overhauser effect experiments for OsFD1 (147–195) S192E were performed in 50 mM potassium phosphate buffer (pH 6.8) containing 50 mM KCl and 1 mM DTT at 283 K, with and without 3 s proton saturation in an interleaved fashion. For backbone chemical shift assignments, the non-uniformly sampled HNCA, HN(CA)CB and HN(COCA)CB spectra were measured.

**Circular dichroism (CD) experiments.** The CD spectrum of OsFD1 (147–195) S192E was recorded between 200 nm and 260 nm in a 0.1-mm cell of a J-720W CD spectropolarimeter (Jasco), with 1.8 mg ml$^{-1}$ protein in 50 mM potassium phosphate buffer (pH 6.8) containing 50 mM KCl and 1 mM DTT, at 293 K. Secondary structure was predicted using Spectra Manager, Windows version 1.0 (Jasco).

**Isothermal titration calorimetry (ITC) experiments.** Before the ITC experiments, purified Hd3a and OsFD1 (S192E) were dialysed against 20 mM MES buffer (pH 6.8) containing 50 mM KCl and 1 mM DTT. ITC experiments were performed at 298 K on a VP-ITC Micro Calorimeter (Microcal). The titration experiments were performed with an initial 5-µl injection, followed by 28 10-µl injections of 0.4 mM Hd3a into 1.8 ml of 30 µM OsFD1 (S192E). The experimental data were analysed with the Origin-ITC software package (OriginLab).

**Crystallization and data collection.** Hd3a (6–170) C43L/C109S/C166S and GF14 (1–235) were used for crystallization. A phosphorylated OsFD1 peptide (187–195, phosphorylated at S192) was chemically synthesized (Operon Biotechnologies). Purified Hd3a and GF14 proteins, at a molar ratio of 1.5:1 and a total concentration of 10 mg ml$^{-1}$, were dialysed against 10 mM Tris-HCl (pH 7.5) and 20 mM NaCl. Crystals of the Hd3a–GF14 complex were grown for 2 weeks at 4 °C by the sitting-drop vapour diffusion method, in a solution containing 0.1 M HEPES (pH 7.5), 0.2 M ammonium sulphate and 25% PEG3350. The Hd3a–GF14–OsFD1 crystals were produced by soaking the Hd3a–GF14 crystals for 15 min in reservoir solution containing 25% ethylene glycol and 2 mM OsFD1 peptide, then flash-cooled in a nitrogen stream at 100 K. Crystallographic data were collected on the BL-5A beamline at the Photon Factory and on the BL41XU beamline at the SPring-8 synchrotron facility. Diffraction data ware processed using the HKL2000 software package[42]. The crystal belongs to space group *P*1 with unit cell parameters $a = 76.7$ Å, $b = 96.6$ Å, $c = 99.5$ Å, $\alpha = 68.2°$, $\beta = 87.9°$, $\gamma = 77.9°$, and contains two Hd3a–GF14–OsFD1 hexamers per asymmetric unit, with a high solvent content of 66.4% (Supplementary Table 1).

**Structure determination and refinement.** The crystal structure was solved by molecular replacement using PHASER software[25]. Two search models were used: tobacco 14-3-3c (ref. 43) (PDB code 2O98) for GF14, and *Arabidopsis* FT[16] (PDB code 1WKP) for Hd3a. Structures were rebuilt in COOT software[25] and refined using REFMAC5 (ref. 26) and CNS[27] software. For Ramachandran plot analysis, Rampage software was used[44]. Final refinement statistics are summarized in Supplementary Table 1. Coordinates of the final model and structure factors are deposited in the Protein Data Bank. All structures in the figures were drawn by PyMOL software[45].

**Gel-shift DNA binding assay.** A C-box element (5′-CTTCACGAGACGTCGATAATCA-3′) and its complementary sequence (5′-TGATTATCGACGTCTCGTGAAG-3′) from the *Arabidopsis AP1* promoter[7], and a region without a C-box (5′- AATGCAAAGCGAAGGTGACACT-3′) and its complementary sequence (5′-AGTGTCACCTTCGCTTTGCATT-3′) from another region of the *Arabidopsis AP1* promoter[7], were used for DNA binding assays. These oligonucleotides were chemically synthesized (Hokkaido System Science), annealed and purified by column chromatography. DNA binding reactions were carried out in 100 mM Tris-borate (pH 8.2), 0.2 mM EDTA, 1.5 mM $MgCl_2$ and 5% glycerol. Each mixture, containing (as appropriate) 30 pmol of C-box DNA, 30 pmol of non-C-box DNA, 40 pmol OsFD1, 40 pmol GF14, 40 pmol Hd3a, 40 pmol OsFD1–GF14, 40 pmol GF14–Hd3a, 40 pmol OsFD1–GF14 with 80 pmol Hd3a, and 40 pmol OsFD1–Hd3a, was incubated for 30 min at 4 °C. The mixtures were separated on a 10% polyacrylamide gel using ×0.5 TBE running buffer and stained with ethidium bromide.

**Model of the FAC–DNA complex.** The model of the FAC–DNA complex shown in Fig. 3b was obtained by superimposing the structure of the mouse CREB bZIP–C-box DNA complex (PDB code 1DH3) onto that of FAC. The CREB bZIP region shows the highest sequence similarity to OsFD1 in PDB, and recognizes C-box DNA[28]. The bZIP length for OsFD1 was predicted by COILS[46], and residues 126–180 were predicted to delineate the coiled-coil region (COILS score >0.9). The amino-acid sequence of OsFD1 residues 126–180 is well conserved in all FD homologues, and is well aligned with CREB bZIP residues 285–339 (55% similarity; no gaps) by ClustalW. Therefore, CREB bZIP residues 285–339 were used as the model of OsFD1 bZIP. The spatial arrangement of GF14 and the helices of OsFD1 bZIP were inferred by comparison with the structure of a complex between tobacco 14-3-3 and $H^+$-ATPase[43]. The OsFD1 bZIP region could be fitted into the wide and deep groove of the Hd3a–GF14 complex without any contacts.


31. Okano, Y., Miki, D. & Shimamoto, K. Small interfering RNA (siRNA) targeting of endogenous promoters induces DNA methylation, but not necessarily gene silencing, in rice. *Plant J.* **53,** 65–77 (2008).
32. Hiei, Y., Ohta, S., Komari, T. & Kumashiro, T. Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *Plant J.* **6,** 271–282 (1994).
33. Kyozuka, J. & Shimamoto, K. *Plant Tissue Culture Manual* (ed. Lindsey, K.) Ch. 1, 1–17 (Kluwer Academic, 1991).
34. Hollenberg, S. M., Sternglanz, R., Cheng, P. F. & Weintraub, H. Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. *Mol. Cell. Biol.* **15,** 3813–3822 (1995).
35. Bartel, P. L., Chien, C., Sternglanz, R. & Fields, S. Using the two-hybrid system to detect protein–protein interactions. In *Cellular Interactions in Development: A Practical Approach* (ed. Hartley, D.) 153–179 (Oxford Univ. Press, 1993).
36. Nakagawa, T. *et al.* Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J. Biosci. Bioeng.* **104,** 34–41 (2007).
37. Hellens, R. P., Edwards, E. A., Leyland, N. R., Bean, S. & Mullineaux, P. M. pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol. Biol.* **42,** 819–832 (2000).
38. The Rowland NMR Toolkit. ⟨http://rnmrtk.uchc.edu/rnmrtk/RNMRTK.html⟩ (2008).
39. Goddard, D. T. & Kneller, G. *SPARKY 3.* (Univ. California, 2004).
40. Rovnyak, D. *et al.* Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction. *J. Magn. Reson.* **170,** 15–21 (2004).
41. Bodenhausen, G. & Ruben, D. J. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69,** 185–189 (1980).


©2011 Macmillan Publishers Limited. All rights reserved

42. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
43. Ottmann, C. *et al.* Structure of a 14-3-3 coordinated hexamer of the plant plasma membrane $H^+$-ATPase by combining X-ray crystallography and electron cryomicroscopy. *Mol. Cell* **25,** 427–440 (2007).
44. Lovell, S. C. *et al.* Structure validation by Cα geometry: Φ, ψ and Cβ deviation. *Proteins* **50,** 437–450 (2003).
45. DeLano, D. L. The PyMOL Molecular Graphics System. ⟨http://www.pymol.org⟩ (2002).
46. Lupas, A., Van Dyke, M. & Stock, J. Predicting coiled coils from protein sequences. *Science* **252,** 1162–1164 (1991).

©2011 **Macmillan Publishers Limited. All rights reserved**

LETTER

doi:10.1038/nature10230

# A forty-kilodalton protein of the inner membrane is the mitochondrial calcium uniporter

Diego De Stefani[1,2]*, Anna Raffaello[1,2]*, Enrico Teardo[3], Ildikò Szabò[3] & Rosario Rizzuto[1,2]

**Mitochondrial $Ca^{2+}$ homeostasis has a key role in the regulation of aerobic metabolism and cell survival[1], but the molecular identity of the $Ca^{2+}$ channel, the mitochondrial calcium uniporter[2], is still unknown. Here we have identified *in silico* a protein (named MCU) that shares tissue distribution with MICU1 (also known as CBARA1), a recently characterized uniporter regulator[3], is present in organisms in which mitochondrial $Ca^{2+}$ uptake was demonstrated and whose sequence includes two transmembrane domains. Short interfering RNA (siRNA) silencing of MCU in HeLa cells markedly reduced mitochondrial $Ca^{2+}$ uptake. MCU overexpression doubled the matrix $Ca^{2+}$ concentration increase evoked by inositol 1,4,5-trisphosphate-generating agonists, thus significantly buffering the cytosolic elevation. The purified MCU protein showed channel activity in planar lipid bilayers, with electrophysiological properties and inhibitor sensitivity of the uniporter. A mutant MCU, in which two negatively charged residues of the putative pore-forming region were replaced, had no channel activity and reduced agonist-dependent matrix $Ca^{2+}$ concentration transients when overexpressed in HeLa cells. Overall, these data demonstrate that the 40-kDa protein identified is the channel responsible for ruthenium-red-sensitive mitochondrial $Ca^{2+}$ uptake, thus providing a molecular basis for this process of utmost physiological and pathological relevance.**

The past two decades have seen the re-appraisal of the key role of mitochondria in decoding the highly pleiotropic $Ca^{2+}$ signals evoked by physiological and pathological stimuli[4]. Mitochondria undergo rapid changes in matrix $Ca^{2+}$ concentration ($[Ca^{2+}]_{mt}$) upon cell stimulation, because their low affinity uptake systems are exposed to microdomains of high $[Ca^{2+}]$ in proximity to endoplasmic reticulum or plasma membrane $Ca^{2+}$ channels[5–9]. In turn, $[Ca^{2+}]_{mt}$ increases upregulate aerobic metabolism[10,11] and sensitize mitochondria to apoptotic challenges, favouring the release of caspase cofactors[12,13]. Mitochondrial $Ca^{2+}$ buffering also shapes the amplitude and spatio-temporal patterns of cytosolic $Ca^{2+}$ concentration ($[Ca^{2+}]_{cyt}$) increases[14–16]. The properties of the $Ca^{2+}$ transporter are still largely those characterized in the 1960s: (1) electrogenic transport (hence the name calcium 'uniporter'), then shown to be a highly selective channel[17], (2) sensitivity to ruthenium red, and (3) low affinity for the cation[1,2,18]. Its molecular identity, however, has remained elusive.

Recently, a uniporter regulator (named MICU1), necessary for rapid mitochondrial $Ca^{2+}$ uptake in intact cells, was identified *in silico* in the MitoCarta database[3]. We thus looked in the same database for the actual transporter, that is, an integral inner membrane protein fulfilling the criteria for being the bona fide MCU. Among 529 candidates with ubiquitous expression in mammalian tissues, we restricted the analysis to 89 proteins with two or more predicted transmembrane domains in the primary sequence. Of these, 20 were absent in *Saccharomyces cerevisiae* (lacking a RuR-sensitive mitochondrial $Ca^{2+}$ uptake route[19]), and, among those, 14 were significantly conserved also in kinetoplastids (*Trypanosoma* and *Leishmania*), in which rapid, uncoupler-sensitive $Ca^{2+}$ fluxes to mitochondria were measured[20,21]. This list (Supplementary Table 1) included 13 proteins with either already characterized functions (complex I subunits, assembly regulators, and so on) or poor alignment scores. The remaining protein (NP_001028431) includes a highly conserved domain comprising two transmembrane regions and an intervening loop enriched in acidic residues (Fig. 1a), indicative of a possible common role in cation permeation. This domain was also detected by hidden Markov model-based analysis of evolutionary conserved proteins (Pfam PF04678 and Panther PTHR1346 domains, with undefined function), further supporting the significance of the hit. We thus considered this protein a putative mitochondrial $Ca^{2+}$ channel, named it MCU (mitochondrial calcium uniporter) and characterized it in detail. Tissue expression profiling by real-time PCR showed ubiquitous presence in all investigated tissues (Fig. 1b), in agreement with a housekeeping role, with expression levels correlating with those of MICU1.

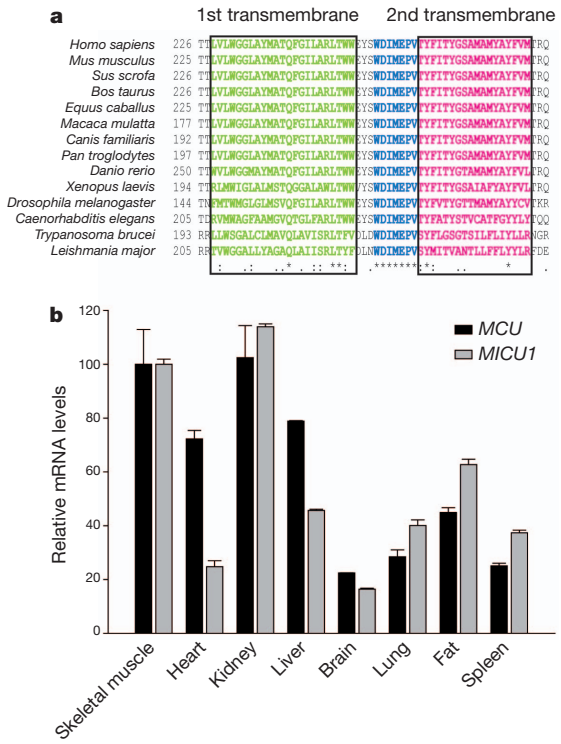
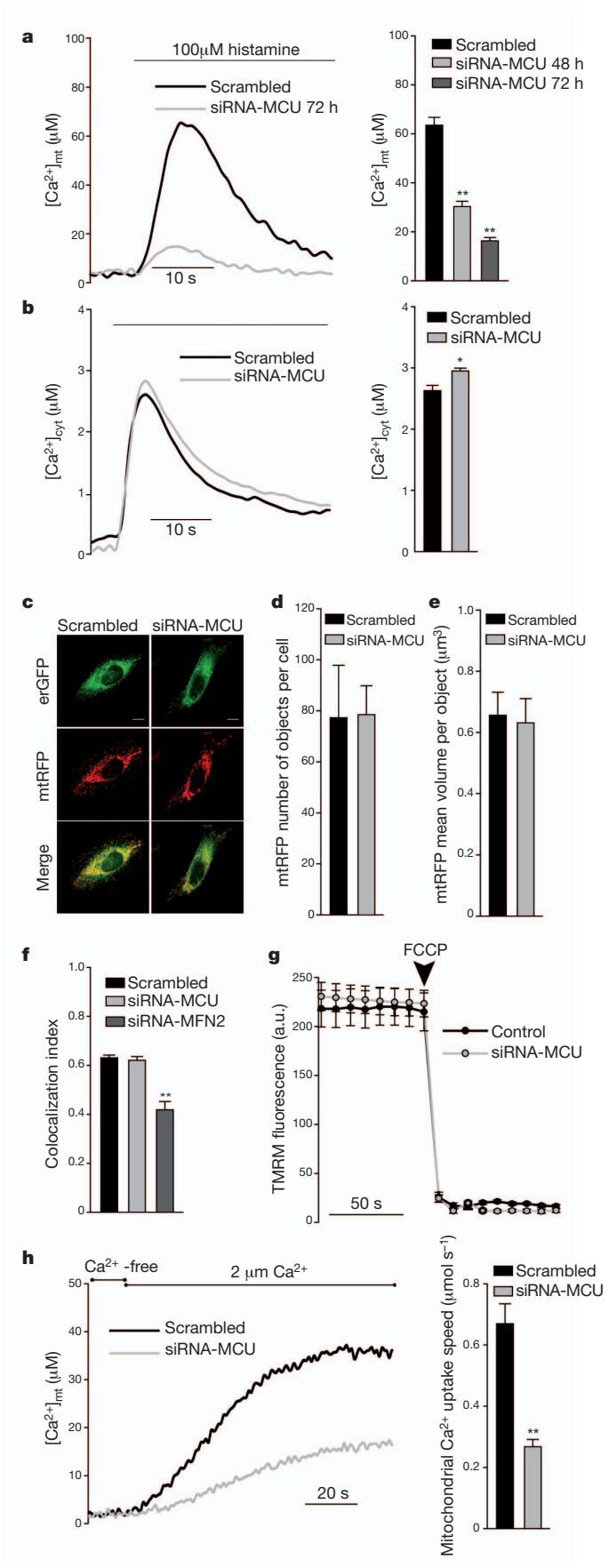



**Figure 1 | MCU includes two highly conserved transmembrane domains and is ubiquitously expressed in mammals, similarly to its putative regulator MICU1. a**, Alignment of the putative transmembrane domain and pore region of MCU proteins from 14 different species . **b**, Real-time PCR analysis of mouse tissues. mRNA extraction and real-time PCR was as described in the Methods section. Expression levels are normalized to skeletal muscle, and presented as means ± s.d. ($n = 3$).

[1]Departments of Biomedical Sciences, University of Padua, 35121 Padua, Italy. [2]CNR Institute of Neuroscience, University of Padua, 35121 Padua, Italy. [3]Department of Biology, University of Padua, 35121 Padua, Italy.
*These authors contributed equally to this work



©2011 Macmillan Publishers Limited. All rights reserved



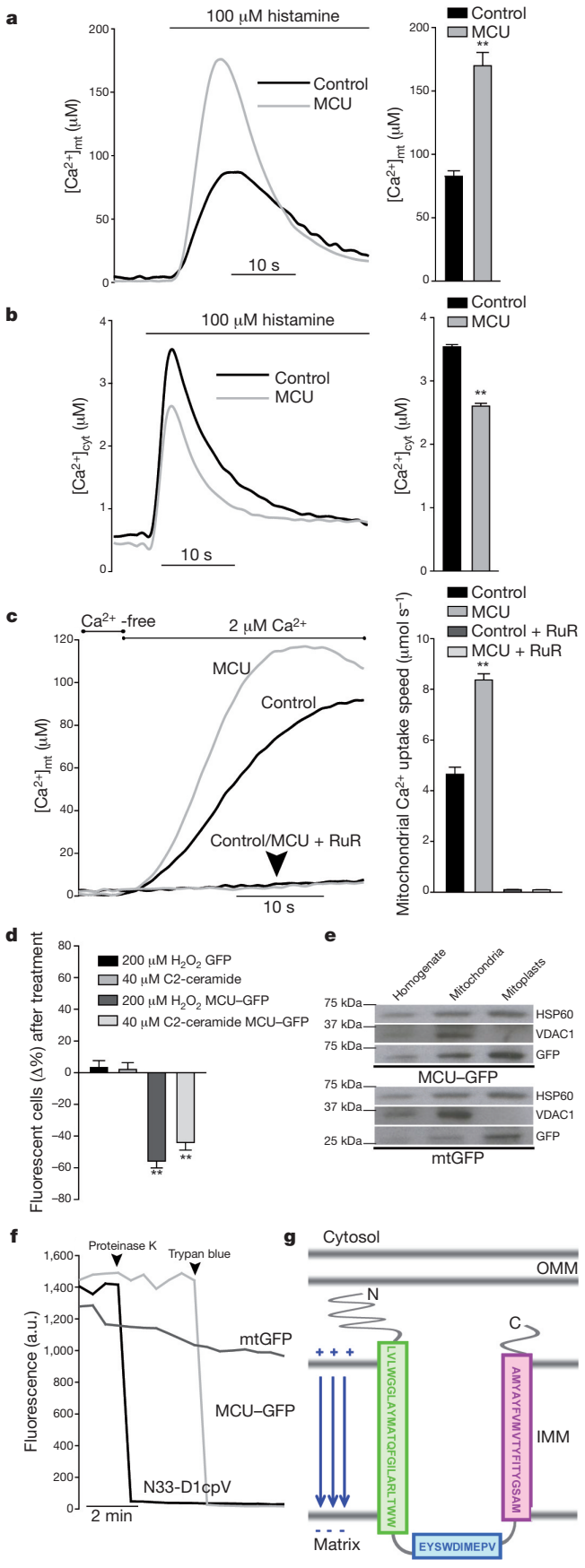

To verify the role of MCU in mitochondrial $Ca^{2+}$ handling, siRNA oligonucleotides were synthesized and tested (Supplementary Fig. 1), $Ca^{2+}$ measurements were then carried out with aequorin-based mitochondrial and cytosolic $Ca^{2+}$ probes (mtAEQ and cytAEQ) (Fig. 2a and Supplementary Fig. 1). MCU-silenced and control HeLa cells, perfused with modified Krebs–Ringer buffer supplemented with 1 mM $CaCl_2$ (KRB), were challenged, where indicated, with 100 μM histamine, an inositol 1,4,5-trisphosphate-generating agonist causing $Ca^{2+}$ release from the endoplasmic reticulum. In control cells, the $[Ca^{2+}]_{mt}$ rise evoked by histamine stimulation was ~63 μM (Fig. 2a). In MCU-silenced cells mitochondrial $Ca^{2+}$ accumulation was markedly reduced (peak value ~30 μM and ~16 μM 48 h and 72 h after silencing, respectively). The marked reduction was specific of mitochondria, as $[Ca^{2+}]_{cyt}$ rises were almost unaffected (rather, a small but significant increase was detected, possibly due to reduced mitochondrial buffering) (Fig. 2b). Overexpression of MICU1 did not rescue the $[Ca^{2+}]_{mt}$ reduction in MCU-silenced cells, and had a marginal effect in control cells (Supplementary Fig. 2). Experiments were then carried out to exclude that the observed effect was secondary to changes in mitochondrial morphology (that is, fragmentation, swelling or redistribution away from the endoplasmic reticulum) or in the driving force for $Ca^{2+}$ accumulation (that is, a collapse of mitochondrial membrane potential, $\Delta\Psi$). On the former aspect, mitochondria and endoplasmic reticulum were labelled in living cells by co-expressing red and green fluorescent protein specifically targeted to the mitochondria and endoplasmic reticulum compartments (mtRFP and erGFP; Fig. 2c and Supplementary Fig. 3). mtRFP imaging showed the typical three-dimensional interconnected network, with no difference in mitochondrial number and volume between control and MCU-silenced cells (Fig. 2d, e). Also the endoplasmic reticulum showed a similar distribution in control and MCU-silenced cells, and no difference was detected in the number of mtRFP/erGFP fluorescence overlaps, whereas in cells in which the endoplasmic reticulum/mitochondria tether mitofusin 2 (*MFN2*)[22] was silenced a marked reduction was observed (Fig. 2f and Supplementary Fig. 3). As to the driving force for $Ca^{2+}$ uptake, no difference was detected in the loading of the fluorescent dye tetramethyl rhodamine methyl ester (TMRM) (Fig. 2g), thus ruling out a $\Delta\Psi$ drop in MCU-silenced cells. In brief, mitochondria simply seem to have an intrinsically reduced ability of accumulating $Ca^{2+}$. Direct evidence was obtained by permeabilizing MCU-silenced and control cells in intracellular buffer, containing 100 μM EGTA ($Ca^{2+}$-free) and then imposing an EGTA-buffered fixed $[Ca^{2+}]$ of 2 μM (2 μM $Ca^{2+}$). Under those conditions, in MCU-silenced cells $Ca^{2+}$ uptake was initiated at a speed that was threefold lower than controls (Fig. 2h).

To demonstrate that MCU promotes $Ca^{2+}$ uptake, overexpression experiments were carried out. The full-length complementary DNA of the protein, obtained by PCR after reverse transcription from skeletal muscle RNA, was identical to that deposited in the NCBI databank (NM_001033259). The cDNA was cloned in pcDNA3.1 and co-expressed with aequorin probes. MCU-overexpressing cells showed a marked increase in the histamine-evoked $[Ca^{2+}]_{mt}$ rise (106%, ~169 μM in MCU-overexpressing vs ~82 μM in control cells, Fig. 3a). The greater mitochondrial response was not secondary to

**Figure 2 | MCU silencing strongly inhibits mitochondrial $Ca^{2+}$ uptake without causing morphological rearrangement or changes in the electrochemical gradient.** **a**, **b**, Measurement of $[Ca^{2+}]_{mt}$ (**a**) and $[Ca^{2+}]_{cyt}$ (**b**) in MCU-silenced cells. **c**, Fluorescence images of mtRFP- and erGFP-labelled mitochondria and endoplasmic reticulum, respectively. **d**, **e**, Mitochondrial number and volume, as deduced by calculating object number (**d**) and size (**e**). **f**, endoplasmic reticulum/mitochondria colocalization, estimated by Pearson's correlation coefficient. **g**, TMRM fluorescence measurements. a.u., arbitrary units. **h**, $[Ca^{2+}]_{mt}$ measurements in permeabilized cells. In this and following figures, experiments are representative of more than five trials, conditions are in the Methods section, and statistics in Supplementary Table 2. $^*P < 0.05$, $^{**}P < 0.001$. Error bars correspond to mean ± s.e.m.



©2011 Macmillan Publishers Limited. All rights reserved




alterations of the cytosolic response. Rather, a significant reduction in the $[Ca^{2+}]_{cyt}$ transient was observed, most probably because of increased $Ca^{2+}$ clearance by mitochondria (Fig. 3b). Enhanced mitochondrial $Ca^{2+}$ uptake was confirmed also upon capacitative $Ca^{2+}$ influx (Supplementary Fig. 4) and in permeabilized cells (Fig. 3c). As in Fig. 2h, mitochondrial $Ca^{2+}$ uptake was initiated by switching the medium, after digitonin permeabilization, from intracellular buffer/EGTA ($Ca^{2+}$-free) to intracellular buffer/$Ca^{2+}$ (3 μM $Ca^{2+}$). MCU-overexpressing cells showed a faster rate of $Ca^{2+}$ uptake and reached a higher plateau level (Fig. 3c). Finally, we verified whether the increased intrinsic ability of mitochondria to accumulate $Ca^{2+}$ correlated with sensitivity to apoptotic challenges (Fig. 3d). Microscopy counts of cell viability after treatment with C2-ceramide or $H_2O_2$ showed that MCU-expressing cells were more efficiently killed, thus confirming the notion that mitochondrial $Ca^{2+}$ loading synergizes with these apoptotic stimuli[12,23].

We then investigated the subcellular distribution of MCU. For this purpose, tagged versions of MCU were generated (MCU–GFP and MCU–Flag), that had the same effect on $[Ca^{2+}]_{mt}$ as MCU (Supplementary Fig. 5). GFP fluorescence and immunocytochemistry labelling with anti-Flag antibody completely overlapped with mitochondrial markers (Supplementary Fig. 6). Then, intra-mitochondrial distribution and topology was investigated. HeLa cells were transfected with MCU–GFP (or the matrix probe mtGFP) and harvested. The mitochondrial fraction was obtained by differential centrifugation and either blotted directly, or subjected to osmotic swelling, thus obtaining the mitoplast fraction devoid of the outer membrane. Both MCU and mtGFP were progressively more enriched in the mitochondrial and mitoplast fractions (Fig. 3e). Mitochondrial localization was also confirmed by immuno-electron microscopy of HeLa cells expressing MCU–Flag (Supplementary Fig. 7). To show the inner membrane localization and get information on the topology of MCU, we carried out an experiment in digitonin-permeabilized cells, in which GFP fluorescence was visualized before and after treatment with proteinase K or Trypan blue, a fluorescence quencher that crosses the outer, but not the inner mitochondrial membrane[9] (Fig. 3f). Proteinase K had no effect on MCU–GFP fluorescence (but abrogated fluorescence of the cytosolic, outer mitochondrial membrane-anchored N33-D1cpV probe, confirming that MCU is not on the outer mitochondrial membrane), whereas Trypan blue abrogated MCU–GFP (but not mtGFP) fluorescence (Fig. 3f), thus demonstrating that MCU is in the inner mitochondrial membrane, with the carboxy terminus most probably located in the intermembrane space (Fig. 3g).

Finally, we investigated the channel activity of the purified protein reconstituted in a planar lipid bilayer. For this purpose, a His-tagged MCU was generated and expressed in two different heterologous expression systems, *Escherichia coli* and wheat-germ cell-free transcription/translation. The purified protein (Supplementary Fig. 8) was inserted into lipid bilayers, and its electrophysiological activity was assessed in a medium containing only $Ca^{2+}$ as cation. With both preparations, we measured channel activity, with the properties previously reported for the uniporter[17,24]: conductance of 6–7 pS (in the

**Figure 3 | MCU overexpression increases mitochondrial $Ca^{2+}$ accumulation in intact and permeabilized cells, buffers cytosolic $[Ca^{2+}]$ rises, and sensitizes to apoptotic stimuli; GFP-tagged MCU demonstrates mitochondrial localization and indicates a putative membrane topology.** **a**, **b**, Measurement of $[Ca^{2+}]_{mt}$ (**a**) and $[Ca^{2+}]_{cyt}$ (**b**) in MCU-silenced cells. **c**, $[Ca^{2+}]_{mt}$ measurements in permeabilized cells. RuR, ruthenium red. **d**, Cell viability upon apoptotic challenge. **e**, Subcellular fractionation of MCU–GFP- and mtGFP-expressing cells and western blotting for GFP, matrix (HSP60) and outer mitochondrial membrane (VDAC1) markers. **f**, GFP proteinase K degradation and fluorescence quenching by Trypan blue in permeabilized cells transfected with mtGFP, MCU–GFP or N33-D1cpV. **g**, Schematic representation of the predicted MCU topology. IMM, inner mitochondrial membrane; OMM, outer mitochondrial membrane. $^{**}P < 0.001$. Error bars: **a**–**c**, mean ± s.e.m.; **d**, **f**, mean ± s.d.



©2011 Macmillan Publishers Limited. All rights reserved

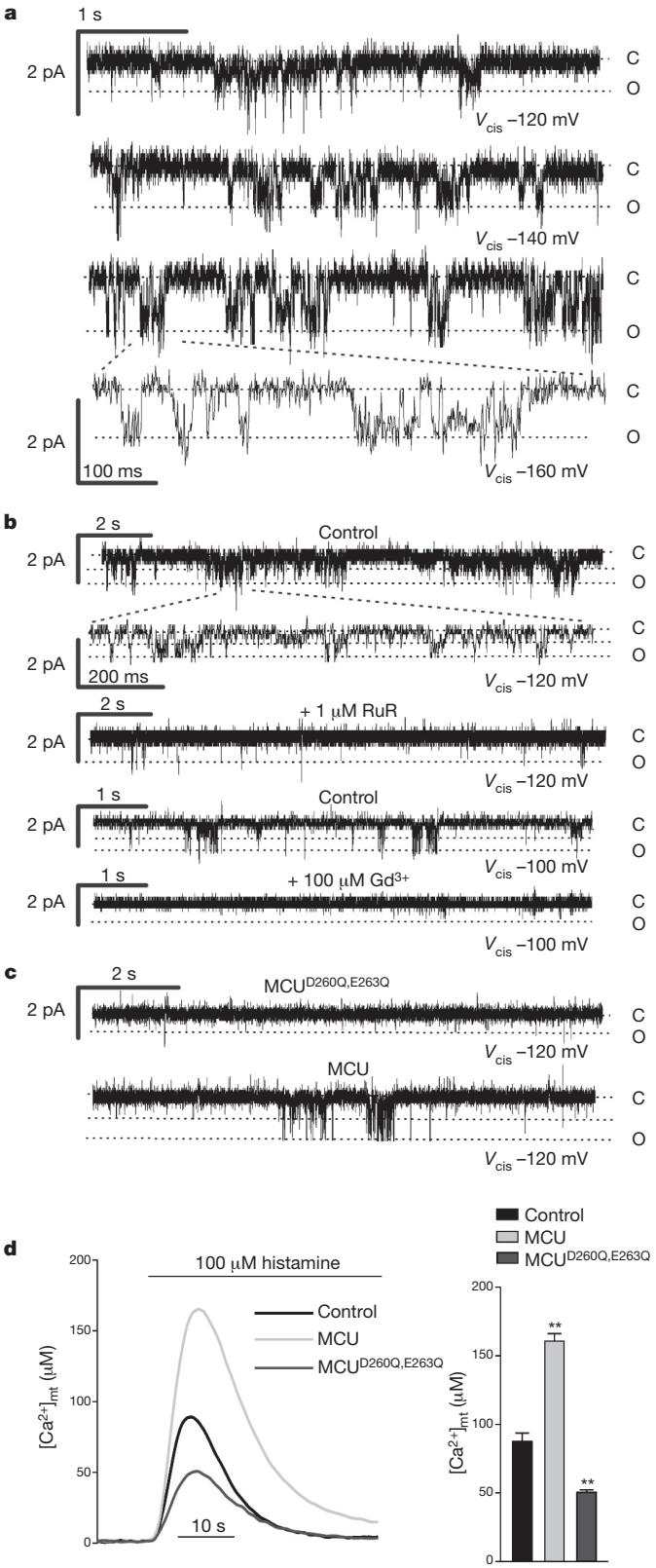



**Figure 4 | Purified MCU shows channel activity in lipid bilayers, whereas $MCU^{D260Q,E263Q}$ shows no channel activity and reduces $[Ca^{2+}]_{mt}$ transients in cells.** **a–c**, Electrophysiological recordings of purified MCU or $MCU^{D260Q,E263Q}$, produced in *E. coli* (**a** and **b**) or *in vitro* (**c**) and reconstituted in planar lipid bilayers. C, closed; O, opened; $V_{cis}$, voltage applied to the *cis* side. **a**, MCU traces at different applied voltages; **b**, MCU traces before and after addition of RuR or $Gd^{3+}$; **c**, $MCU^{D260Q,E263Q}$ traces. After 10 min recording with no current, MCU was added and channel activity was detected after 1–2 min (lower trace). **d**, $[Ca^{2+}]_{mt}$ measurements in $MCU^{D260Q,E263Q}$-expressing cells. Error bars correspond to mean ± s.e.m.

negative voltage range applied in the *cis* compartment), fast kinetics, low probability of opening at low voltages ($P_o$), $P_o$ increase with voltage, and inhibition by ruthenium red and the lanthanide gadolinium ($Gd^{3+}$) (Fig. 4a and b and Supplementary Fig. 9). Then, to get evidence of cation permeation across the putative pore-forming domain, we mutated into glutamines two negatively charged residues of the conserved region (D260Q and E263Q) and assessed the electrophysiological properties of the mutant protein ($MCU^{D260Q,E263Q}$) produced *in vitro*. $MCU^{D260Q,E263Q}$ failed to give rise to $Ca^{2+}$-permeable channel activity in bilayer experiments, whereas in the same membrane the subsequent addition of MCU initiated channel activity with the characteristics described above (Fig. 4c). $MCU^{D260Q,E263Q}$, when overexpressed in HeLa cells, showed the same membrane topology as MCU (Supplementary Fig. 10) and caused a marked reduction in histamine-induced $[Ca^{2+}]_{mt}$ rises, both compared to MCU-overexpressing and control cells (Fig. 4d). $MCU^{D260Q,E263Q}$ thus acts as a dominant-negative mutant, most probably by either inhibiting MCU channel activity within an oligomer or competing for docking sites or essential regulators.

In conclusion, we have identified a 40-kDa protein which fulfils the requirements for being the long searched for mitochondrial calcium uniporter: it contains two transmembrane domains and shows channel activity *in vitro* with the previously characterized properties of the uniporter, it is localized in the inner membrane and, finally, it markedly enhances mitochondrial $Ca^{2+}$ uptake upon overexpression. The availability of molecular information opens the mitochondrial checkpoint in cellular $Ca^{2+}$ signalling to a deeper understanding. Informative animal models can now be generated and new drugs developed to influence processes regulated by mitochondrial $Ca^{2+}$ signals, such as aerobic metabolism and cell death.

## METHODS SUMMARY

Candidate mouse genes coding for the mitochondrial calcium uniporter were searched starting from the MitoCarta database. Bioinformatic analysis was performed by selecting candidates according to the following parameters: (1) genes expressed in most mammalian tissues, (2) the presence of two or more transmembrane domains, (3) without any homologues in *Saccharomyces cerevisiae* but (4) conserved in kinetoplastids. The resulting 14 candidates were manually screened to search for highly conserved putative channel-forming domains.

All the experiments were performed in HeLa cells. MCU was silenced using specific siRNAs. The MCU cDNA was retro-transcribed and amplified from mouse skeletal muscle messenger RNA and overexpressed after cloning in several expression plasmids.

$[Ca^{2+}]_{cyt}$ and $[Ca^{2+}]_{mt}$ were measured in MCU-silenced and MCU-overexpressing cells using appropriate cytosolic and mitochondrial probes[25].

Morphology of MCU-silenced cells was evaluated using GFP and RFP specifically targeted to the mitochondria and endoplasmic reticulum compartments (erGFP and mtRFP)[6].

MCU mitochondrial localization was demonstrated by confocal imaging of cells transfected with the MCU–GFP expression vector and loaded with MitoTracker Red and by fractionation and western blotting of cells overexpressing MCU–GFP, as previously described[26].

Reconstitution in planar lipid bilayer was carried out with the purified protein from both *E. coli* and *in vitro* transcription/translation using 100 mM $CaCl_2$. Bilayers of approximately 150–200 pF capacity were prepared using purified soybean azolectin.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 29 December 2010; accepted 27 May 2011.**
**Published online 19 June 2011.**

1. Szabadkai, G. & Duchen, M. R. Mitochondria: the hub of cellular $Ca^{2+}$ signaling. *Physiology (Bethesda)* **23,** 84–94 (2008).
2. Carafoli, E. Historical review: mitochondria and calcium: ups and downs of an unusual relationship. *Trends Biochem. Sci.* **28,** 175–181 (2003).



©2011 Macmillan Publishers Limited. All rights reserved

3. Perocchi, F. *et al. MICU1* encodes a mitochondrial EF hand protein required for $Ca^{2+}$ uptake. *Nature* **467,** 291–296 (2010).
4. Berridge, M. J., Bootman, M. D. & Roderick, H. L. Calcium signalling: dynamics, homeostasis and remodelling. *Nature Rev. Mol. Cell Biol.* **4,** 517–529 (2003).
5. Rizzuto, R., Brini, M., Murgia, M. & Pozzan, T. Microdomains with high $Ca^{2+}$ close to IP3-sensitive channels that are sensed by neighboring mitochondria. *Science* **262,** 744–747 (1993).
6. Rizzuto, R. *et al.* Close contacts with the endoplasmic reticulum as determinants of mitochondrial $Ca^{2+}$ responses. *Science* **280,** 1763–1766 (1998).
7. Csordás, G., Thomas, A. P. & Hajnóczky, G. Quasi-synaptic calcium signal transmission between endoplasmic reticulum and mitochondria. *EMBO J.* **18,** 96–108 (1999).
8. Csordás, G. *et al.* Imaging interorganelle contacts and local calcium dynamics at the ER-mitochondrial interface. *Mol. Cell* **39,** 121–132 (2010).
9. Giacomello, M. *et al.* $Ca^{2+}$ hot spots on the mitochondrial surface are generated by $Ca^{2+}$ mobilization from stores, but not by activation of store-operated $Ca^{2+}$ channels. *Mol. Cell* **38,** 280–290 (2010).
10. Hajnóczky, G., Robb-Gaspers, L. D., Seitz, M. B. & Thomas, A. P. Decoding of cytosolic calcium oscillations in the mitochondria. *Cell* **82,** 415–424 (1995).
11. Jouaville, L. S., Pinton, P., Bastianutto, C., Rutter, G. A. & Rizzuto, R. Regulation of mitochondrial ATP synthesis by calcium: evidence for a long-term metabolic priming. *Proc. Natl Acad. Sci. USA* **96,** 13807–13812 (1999).
12. Pinton, P. *et al.* The $Ca^{2+}$ concentration of the endoplasmic reticulum is a key determinant of ceramide-induced apoptosis: significance for the molecular mechanism of Bcl-2 action. *EMBO J.* **20,** 2690–2701 (2001).
13. Pacher, P. & Hajnóczky, G. Propagation of the apoptotic signal by mitochondrial waves. *EMBO J.* **20,** 4107–4121 (2001).
14. Hajnóczky, G., Hager, R. & Thomas, A. P. Mitochondria suppress local feedback activation of inositol 1,4,5-trisphosphate receptors by $Ca^{2+}$. *J. Biol. Chem.* **274,** 14157–14162 (1999).
15. Boitier, E., Rea, R. & Duchen, M. R. Mitochondria exert a negative feedback on the propagation of intracellular $Ca^{2+}$ waves in rat cortical astrocytes. *J. Cell Biol.* **145,** 795–808 (1999).
16. Tinel, H. *et al.* Active mitochondria surrounding the pancreatic acinar granule region prevent spreading of inositol trisphosphate-evoked local cytosolic $Ca^{2+}$ signals. *EMBO J.* **18,** 4999–5008 (1999).
17. Kirichok, Y., Krapivinsky, G. & Clapham, D. E. The mitochondrial calcium uniporter is a highly selective ion channel. *Nature* **427,** 360–364 (2004).
18. Nicholls, D. G. Mitochondria and calcium signaling. *Cell Calcium* **38,** 311–317 (2005).
19. Uribe, S., Rangel, P. & Pardo, J. P. Interactions of calcium with yeast mitochondria. *Cell Calcium* **13,** 211–217 (1992).
20. Xiong, Z. H., Ridgley, E. L., Enis, D., Olness, F. & Ruben, L. Selective transfer of calcium from an acidic compartment to the mitochondrion of *Trypanosoma brucei*. Measurements with targeted aequorins. *J. Biol. Chem.* **272,** 31022–31028 (1997).
21. Benaim, G., Bermudez, R. & Urbina, J. A. $Ca^{2+}$ transport in isolated mitochondrial vesicles from *Leishmania braziliensis* promastigotes. *Mol. Biochem. Parasitol.* **39,** 61–68 (1990).
22. de Brito, O. M. & Scorrano, L. Mitofusin 2 tethers endoplasmic reticulum to mitochondria. *Nature* **456,** 605–610 (2008).
23. Hajnóczky, G., Csordás, G., Madesh, M. & Pacher, P. Control of apoptosis by $IP_3$ and ryanodine receptor driven calcium signals. *Cell Calcium* **28,** 349–363 (2000).
24. Michels, G. *et al.* Regulation of the human cardiac mitochondrial $Ca^{2+}$ uptake by 2 different voltage-gated $Ca^{2+}$ channels. *Circulation* **119,** 2435–2443 (2009).
25. Pinton, P., Rimessi, A., Romagnoli, A., Prandini, A. & Rizzuto, R. Biosensors for the detection of calcium and pH. *Methods Cell Biol.* **80,** 297–325 (2007).
26. Wieckowski, M. R., Giorgi, C., Lebiedzinska, M., Duszynski, J. & Pinton, P. Isolation of mitochondria-associated membranes and mitochondria from animal tissues and cells. *Nature Protocols* **4,** 1582–1590 (2009).


**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank M. Zoratti for help in the analysis of electrophysiological data, G. Merli for carrying out some of the experiments and P. Bernardi, T. Pozzan and L. Cendron for helpful discussions and for the N33-D1cpV expression plasmid (T. Pozzan ). This research was supported by grants from the Italian Ministry of Education, University and Research, European Commission (FP7 ''MyoAGE'', no. 223576), National Institutes of Health (Grant #1P01AG025532-01A1), Cariparo Foundation (Padua), the Italian Association for Cancer Research (AIRC) and Telethon-Italy (GPP1005).

**Author Contributions** D.D.S. performed bioinformatic analysis, $Ca^{2+}$ measurements and morphological analysis of organelles. A.R. performed molecular biology and gene expression analysis. D.D.S. and A.R. contributed equally to the study. E.T. expressed and purified the protein in heterologous systems; I.S. performed and analysed the electrophysiology experiments; R.R. discussed the results and wrote the paper.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.R. (rosario.rizzuto@unipd.it).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Bioinformatic screening.** The MitoCarta database was screened for ubiquitously expressed proteins (expressed in at least 12 out of 14 tissues analysed, obtaining 529 candidates) and proteins containing two or more transmembrane domains were selected using the TMHMM algorithm[27]. Out of 1,098 proteins present in MitoCarta, 89 fulfilled these requirements. Sequences of these candidates were retrieved and aligned with the *Saccharomyces cerevisiae* proteome by a standard Gonnet matrix to exclude homologues. All alignments with a $P$-value $< 0.001$ were excluded. Out of 89, 20 remained and were subsequently aligned with kinetoplastids proteomes using the TriTrypDB server. Out of 20, 14 had significant scores. These candidates were manually aligned through different species to look for putative highly conserved channel-forming domains.

**RNA extraction and gene expression analyses.** For the expression analysis of MCU and MICU1 in mouse tissues, adult male C57BL/6 mice (28–30 g) were used. Skeletal muscles (tibialis anterior), heart, brain, spleen, lung, liver, kidney and visceral fat were excised from three age-matched animals. Total RNA was purified using an SV Total RNA Isolation kit (Promega) following manufacturer instructions. The RNA was quantified and controlled for its quality using a RNA 6000 LabChip kit (Agilent Technologies) in conjunction with an Agilent Bioanalyzer 2001. An equal amount of total RNA from the three animals was pooled together for each organ, complementary DNA was generated with a cDNA synthesis kit (SuperScript VILO, Invitrogen) and analysed by real-time PCR using the SYBR green chemistry (Bio-Rad). The primers were designed and analysed with Primer3. Identity of the amplicons was confirmed by their dissociation profiles and gel analysis. Real-time PCR standard curves were constructed by using serial dilutions of pooled cDNAs of the analysed samples, using at least four dilution points and the efficiency of all primer sets was between 95 and 105%. Real-time PCR primer sequences were as follows: MICU1-fw, 5′-GTCGAACTCTCGGACCATGT-3′; MICU1-rv, 5′-CAAAGTCCCAGGCAGTTTCT-3′. These primers amplify a fragment of 199 base pairs. MCU-fw, 5′-AAAGGAGCCAAAAAGTCACG-3′; MCU-rv, 5′-AACGGCGTGAGTTACAAACA-3′. These primers amplify a fragment of 200 bp.

For the expression analysis of MCU silencing, total RNA was purified from HeLa cells transfected with scrambled, siRNA-MCU#1 and siRNA-MCU#2 for 48 h following the standard TRIzol protocol. The RNA was quantified, controlled for its quality and retrotranscribed as described earlier. Complementary DNA was analysed by real-time PCR using the SYBR Green chemistry (Bio-Rad). All data were normalized to *GAPDH* expression. The oligonucleotide primers specific for MCU are the same as above. The oligonucleotide primers specific for GAPDH are the following: GADPH-fw, 5′-CACCATCTTCCAGGAGCGAG-3′; GADPH-rv, 5′-CCTTCTCCATGGTGGTGAAGAC-3′. These primers amplify a fragment of 101 bp.

**Constructs and siRNA.** Mouse MCU (NM_001033259) was amplified from mouse skeletal muscle cDNA by PCR using the following primers.

For cloning into pEGFP-N1: fw, 5′-GAATTCGCCACCATGGCGGCCGCCGCAGGTAG-3′; rv, 5′-GGATCCACTTCCTTTTCTCCGATCTGTCG-3′. The PCR fragment was cloned into EcoRI and BamHI sites in pEGFP-N1 (Clontech).

For cloning into pcDNA3.1: fw, 5′-GGTACCGCCACCATGGCGGCCGCCGCAGGTAG-3′; rv, 5′-GAATTCTCATTCCTTTTCTCCGATCTGTC-3′. The PCR fragment was cloned into KpnI and EcoRI sites in pcDNA3.1 (Invitrogen).

For the cloning of MCU–Flag into pcDNA3.1: fw, 5′-GGTACCGCCACCATGGCGGCCGCCGCAGGTAG-3′; rv, 5′-GGAATTCTCACTTATCGTCGTCATCCTTGTAATCTTCCTTTTCTCCGATCTGTC-3′. The PCR fragment was cloned into KpnI and EcoRI sites in pcDNA3.1 (Invitrogen).

For the cloning of MCU in pET-28A(+): fw, 5′-AGGATCCATGGCGGCCGCCGCAGGTAG-3′; rv, 5′-ACTCGAGTCATTCCTTTTCTCCGATCT-3′. The PCR fragment was cloned into BamHI and XhoI sites in pET-28A(+) (Novagen).

For the cloning of MCU deleted of the mitochondrial targeting sequence (amino acids 1–54) ($MCU^{\Delta MTS}$) in pET-28A(+): fw, 5′-CATATGGCTTCCTGGCAGAGCGTGGG-3′; rv, 5′-CTCGAGTCATTCCTTTTCTCCGATCT-3′. The PCR fragment was cloned into NdeI and XhoI sites in pET-28A(+) (Novagen).

For the cloning of MCU in pIVEX 1.3 WG: fw, 5′-ACATATGGCGGCCGCCGCAGGTAGATC-3′; rv, 5′-TCTCGAGTTCCTTTTCTCCGATCTGTC-3′. The PCR fragment was cloned into NdeI and XhoI sites in pIVEX 1.3 WG (Roche).

The generation of the pcDNA3.1-$MCU^{D260Q,E263Q}$-Flag and pEGFP-N1-$MCU^{D260Q,E263Q}$ mutant was performed by mutagenesis PCR using the wild type pcDNA3.1-MCU-Flag and pEGFP-N1-MCU vectors as template and the mutagenesis primer: 5′-CTGGTGGGAGTACTCGTGG**C**A**A**ATCATG**C**A**A**CCCGTCACCTACTTCATCAC-3′.

Mouse MICU1 (NM_144822) was amplified from mouse skeletal muscle cDNA by PCR using the following primers:

For the cloning of MICU1-HA in pcDNA3.1: fw, 5′-CGGATCCGCCACCATGTTTCGTCTTAACACCCT-3′; rv, 3′-GCTCGAGTCACAGGGAAGCGTAGTCAGGCACATCGTAGGGGTATTTGGGCAGAGCAAAGTCCC-5′. The PCR fragment was cloned into BamHI and XhoI sites in pcDNA3.1 (Invitrogen).

For the cloning of MICU1 in pEGFP-N1: fw, 5′-CCTCGAGATGTTTCGTCTTAACACCCT-3′; rv, 5′-CGGATCCCGTTTGGGCAGAGCAAAGTCCC-3′. The PCR fragment was cloned into XhoI and BamHI sites in pEGFP-N1 (Clontech).

To silence MCU specific siRNA were designed: siRNA-MCU#1: nucleotides 899–917 of the corresponding mRNA (5′-GCCAGAGACAGACAAUACUtt-3′ and 3′-ttCGGUCUCUGUCUGUUAUGA-5′. siRNA-MCU#2: nucleotides 360–378 of the corresponding mRNA (5′-GGGAAUUGACAGAGUUGCUtt-3′and 3′-ttCCCUUAACUGUCUCAACGA-5′). The non-targeting siRNA (scrambled) is the following: 5′-GCCUAAGAACGACAAAUCAtt-3′ and 3′-ttCGGAUUCUUGCUGUUUAGU-5′.

**Cell culture and transfection.** In all the experiments HeLa cells were used. Cells were grown in Dulbecco's modified Eagle's medium (DMEM) (Euroclone), supplemented with 10% fetal bovine serum (FBS) (Euroclone) and transfected with a standard calcium-phosphate procedure. For aequorin measurements, the cells were seeded 24 h before transfection onto 13 mm glass coverslips and allowed to grow to 50% confluence before transfection. For morphologic analyses cells were seeded 24 h before transfection onto 24-mm glass coverslips and allowed to grow to 50% confluence before transfection, unless otherwise specified.

**Aequorin measurements.** HeLa cells grown on 13-mm round glass coverslips at 50% confluence were transfected with the cytosolic (cytAEQ) or the low-affinity mitochondrial (mtAEQmut, referred in the text as mtAEQ) probe (as previously described[25]) together with the indicated siRNA or plasmid. pcDNA3.1 was used as control unless otherwise indicated. The coverslip with the cells was incubated with 5 μM coelenterazine for 1–2 h in KRB (Krebs–Ringer modified buffer: 125 mM NaCl, 5 mM KCl, 1 mM $Na_3PO_4$, 1 mM $MgSO_4$, 5.5 mM glucose, 20 mM HEPES, pH 7.4, 37 °C) supplemented with 1 mM $CaCl_2$, and then transferred to the perfusion chamber. All aequorin measurements were carried out in KRB. Agonists and other drugs were added to the same medium, as specified in the text. The experiments were terminated by lysing the cells with 100 μM digitonin in a hypotonic $Ca^{2+}$-rich solution (10 mM $CaCl_2$ in $H_2O$), thus discharging the remaining aequorin pool. The light signal was collected and calibrated into $[Ca^{2+}]$ values by an algorithm based on the $Ca^{2+}$ response curve of aequorin at physiological conditions of pH, $[Mg^{2+}]$ and ionic strength, as previously described[25]. Representative traces are shown in the figures whereas the full data set is included in Supplementary Table 2. In the experiments with permeabilized cells, a buffer mimicking the cytosolic ionic composition, (intracellular buffer) was used: 130 mM KCl, 10 mM NaCl, 2 mM $K_2HPO_4$, 5 mM succinic acid, 5 mM malic acid, 1 mM $MgCl_2$, 20 mM HEPES, 1 mM pyruvate, 0.5 mM ATP and 0.1 mM ADP (pH 7 at 37 °C). Intracellular buffer was supplemented with either 100 μM EGTA (intracellular buffer/EGTA) or a 2 mM EGTA and 2 mM HEEDTA-buffered $[Ca^{2+}]$ of 2 or 3 μM (intracellular buffer/$Ca^{2+}$), calculated with the Chelator software[28]. HeLa cells were permeabilized by a 1-min perfusion with 50 μM digitonin (added to intracellular buffer/EGTA) during luminescence measurements. Mitochondrial $Ca^{2+}$ uptake speed was calculated as the first derivative by using the SLOPE Excel function (Microsoft) and smoothed for three time points. The higher value reached during $Ca^{2+}$ addition represents the maximal $Ca^{2+}$ uptake speed. $[Ca^{2+}]_{mt}$ following capacitative $Ca^{2+}$ influx was measured by perfusing HeLa cells with the SERCA blocker cyclopiazonic acid (CPA, 20 μM) in a KRB solution containing no $Ca^{2+}$ and 100 μM EGTA. In this protocol, mitochondrial $Ca^{2+}$ uptake were evoked by adding 2 mM $CaCl_2$ to the medium. All of the results are expressed as means ± s.e.m., and Student's $t$-test was used for the statistics. All the materials were from Sigma Aldrich unless specified.

**Mitochondrial and endoplasmic reticulum morphology analysis.** Images for investigating mitochondrial morphology and endoplasmic reticulum/mitochondrial contact sites were taken on a Leica TCS-SP5-II equipped with a PlanApo 100×/1.4 numerical aperture objective. For all images, pinhole was set to 1 Airy unit, pixel size was 75 nm and a $Z$-stack was acquired for the whole depth of the cell by sampling at 130 nm in the $Z$ plane. An argon laser line (488 nm) was used to excite GFP or AlexaFluor488 and the signal was collected in the 492–537 nm range, whereas RFP (and MitoTracker Red and AlexaFluor546) fluorescence was excited by the 543 nm HeNe laser and its emission was collected in the 555–700 nm range. For each image, PMT gain was slightly adjusted in order to maximize signal and avoid saturation. Images were then all analysed with the Fiji image processing package based on ImageJ. Mitochondrial morphology was calculated with the 3D Object Counter plugin, and colocalization was evaluated with the JaCoP plugin[29]. The colocalization index is represented by Pearson's coefficient calculated following Costes randomization (100 cycles) and automatic threshold calculation[30].

**Immunofluorescence.** HeLa cells were grown on 24-mm coverslips and transfected with MCU–Flag encoding plasmid when 50% confluent. After 24 h, cells were washed with PBS, fixed in 4% formaldehyde for 10 min and quenched with 50 mM $NH_4Cl$ in PBS. Cells were permeabilized for 10 min with 0.1% Triton X-100 in PBS and blocked in PBS containing 2% BSA and 0.05% Triton X-100

©2011 Macmillan Publishers Limited. All rights reserved

for 1 h. Cells were then incubated with primary antibodies (anti-HSP60 and anti-Flag) for 3 h at room temperature and washed three times with 0.1% Triton X-100 in PBS. The appropriate isotype-matched, AlexaFluor-conjugated secondary antibodies (Invitrogen) were used and coverslips were mounted with ProLong Gold Antifade reagent (Invitrogen). Images were taken on a Leica SP5-II as detailed earlier.

**Fractionation and western blotting.** HeLa cells ($10^8$) were homogenized with a Dounce homogenizer (100 strokes) and centrifuged at 600*g* to remove entire cells. The supernatant was then centrifuged at 8,000*g* to pellet the crude mitochondrial fraction. Mitoplasts were obtained by osmotic swelling by incubating the mitochondrial fraction in 20 mM HEPES for 20 min. Proteins were quantified using the BCA Protein Assay Kit (Pierce) following the manufacturer instructions. Proteins (10 μg) were separated by SDS–PAGE, transferred onto nitrocellulose membranes (GE Healthcare) and probed using the following antibodies: anti-VDAC1(Abnova), anti-HSP60 (Santa Cruz) and anti-GFP (Abcam). Isotype-matched, horseradish-peroxidase-conjugated secondary antibodies (Bio-Rad) were used, followed by detection by chemiluminescence (GE Healthcare).

To test the siRNAs efficacy, HeLa cells were grown on 10-cm Petri dishes and transfected when 30% confluent. MCU–Flag encoding plasmid (8 μg) was co-transfected with scrambled, siRNA-MCU#1 or siRNA-MCU#2. After 72 h cells were washed, harvested and lysed in RIPA buffer (150 mM NaCl, 25 mM TRIS, 1% Triton-X100, 0.5% Na-deoxycholate and 0.1% SDS, pH 8) added with proteases and phosphatases inhibitor cocktails (Roche). Proteins were quantified by the bicinchoninic acid assay (BCA) method and 25 μg of each sample were loaded on a Novex NuPage Bis-Tris 4–12% precast gel (Invitrogen), transferred onto nitrocellulose membranes and probed with anti-β-tubulin (Sigma) and anti-Flag (Sigma) antibodies. Isotype-matched, horseradish-peroxidase-conjugated secondary antibodies (Bio-Rad) were used, followed by detection by chemiluminescence (GE Healthcare).

**Measurements of mitochondrial $\Delta\Psi$.** Mitochondrial $\Delta\Psi$ was measured by loading cells with 20 nM tetramethyl rhodamine methyl ester (TMRM, Invitrogen) for 30 min at 37 °C. Images were taken on a inverted microscope (Zeiss Axiovert 100) equipped with a PlanFluar 40×/1.3 numerical aperture objective, a Photometrics Coolsnap ES and a LED-based illumination system (Cairn OptoLed II). TMRM excitation was performed at 560 nm and emission was collected through a 590–650 nm band-pass filter. Images were taken every 10 s with a fixed 200 ms exposure time. FCCP (carbonyl cyanide p-trifluoromethoxyphenylhydrazone, 10 μM), an uncoupler of oxidative phosphorylation, was added after 12 acquisitions to completely collapse the electrical gradient established by the respiratory chain ($\Delta\Psi$).

**Trypan blue quenching experiment.** After transfection with mtGFP, MCU–GFP or N33-D1cpV, cells were analysed on spinning disk imaging system composed of an inverted microscope (Zeiss Axiovert 200) equipped with a PlanApo 63×/1.4 numerical aperture objective, a BD-CARVII confocal head, a Photometrics Cascade 512B camera, a piezoelectric *Z* motor (Phisik instruments) and a 300 W Xenon bulb (Sutter Instruments). At each time point, a *Z*-stack was taken (*Z*-step: 500 nm) using standard EGFP filter-set (Semrock GFP-A-Basic-000) and the best focus plane was selected for subsequent analysis. After 1 min permeabilization in intracellular buffer containing 50 μM digitonin, cells were incubated in 1 ml of intracellular buffer, first treated with 4 units of proteinase K and then with 0.005% trypan blue. Best focus images of each time point were background-corrected and mitochondrial mean fluorescence intensity was measured with the MetaMorph software (Molecular devices).

**Cell death experiment.** Cell sensitivity to apoptotic stimuli was evaluated as previously described[12]. HeLa cells grown on 24-mm round glass coverslips at 30% confluence were transfected with MCU–GFP or GFP alone. The effect on cell fate was evaluated by applying an apoptotic challenge (40 μM C2-ceramide or 200 μM $H_2O_2$) and comparing the survival of transfected and non-transfected cells. In these experiments, the percentage of GFP-positive cells was calculated before and after applying an apoptotic stimulus (C2-ceramide or $H_2O_2$). In mock-transfected cells, although the total number of cells is reduced after cell death induction, the apparent transfection efficiency was maintained (that is, transfected and non-transfected cells have the same sensitivity to the apoptotic stimulus and thus die to the same extent). However, when cells are transfected with a construct influencing their sensitivity to apoptosis, this will be reflected by a change in the fraction of fluorescent cells, that is, in the 'apparent' transfection efficiency. Thus, protection from apoptosis results into an apparent increase of transfection, whereas a decrease reflects a higher sensitivity to apoptosis. Data are reported as the mean percentage change in the apparent transfection efficiency after apoptotic challenge compared to vehicle-treated cells. Cells were extensively washed with PBS, stained with 4′,6-diamidino-2-phenylindole (DAPI) and two images per field (blue and green fluorescence) were taken at 20× magnification (mean transfection efficiencies were roughly 30% for both GFP and MCU–GFP). At least 10 fields per coverslip were randomly imaged and counted. Data presented are the sum of at least two different coverslip per experimental condition carried out in three different independent experiments.

**Immunogold electron microscopy.** HeLa cells were transfected with control (pcDNA3.1) or MCU–Flag plasmids. Cells were fixed in 4% formaldehyde and 0.1% glutaraldehyde in 0.1 M cacodylate buffer 24 h after transfection. Dehydration was carried out by subsequent incubation with increasing ethanol concentrations (20, 40, 60, 80 and 100%). Cells were then infiltrated and embedded in LRW resin and 100-nm slices were obtained. For immunogold staining slices were first blocked in PBS containing 2% BSA and 0.05% Triton X-100 and then incubated for 1 h with anti-Flag (Sigma) antibody. After five washes with PBS, slices were incubated for 1 h with an isotype matched 10-nm gold-particles-conjugated secondary antibody (Sigma). After five washes with PBS and one wash with $H_2O$, slices were contrasted with uranyl acetate for 15 min and lead citrate for 6 min. Slices observation and imaging were carried out on a Tecnai $G^2$ TEM.

**Protein expression and purification.** *E. coli*: MCU and $MCU^{\Delta MTS}$ were cloned into the pET-28A(+) vector, as described above. Competent cells were transformed with the indicated constructs. Induction was performed at 20 °C ($D_{600} = 0.4$) with 0.35 mM isopropyl-β-D-thiogalactoside (IPTG) for 24 h. After sonication in buffer A (50 mM Na-phosphate, pH 7.4, 300 mM NaCl and protease inhibitor cocktail), the membranous fraction was collected by centrifugation (10 min at 12000*g*) and solubilized with *n*-decyl-β-D-maltopyranoside (2.5%) for 3 h at room temperature. The non-solubilized material was removed by centrifugation and the supernatant was loaded on a nickel chromatography column (HIS-Select Nickel affinity gel, Sigma). Column was washed with buffer A and the protein was eluted with imidazole gradient (50–200 mM). Fractions of 500 μl were collected and dialysed against buffer A without NaCl. A concentration of 0.025% of detergent was maintained throughout the purification.

*In vitro*: MCU was cloned into pIVEX 1.3 WG, as described above. *In vitro* expression was performed by using a RTS100 Wheat Germ CECF Kit (Roche). After expression, the reaction mix was solubilized with 2% Triton X-100 for 90 min at 30 °C under shaking.

Gel electrophoresis: SDS–PAGE was performed using 6 M urea and standard protocols. Each lane was loaded with 30 μl of eluted fractions and 1 μl of the reaction mix.

**Electrophysiology.** Electrophysiology experiments were carried out as previously described[31,32]. A Warner Instruments (Hamden) electrophysiological planar bilayer apparatus was used. Bilayers of approximately 150–200 pF capacity were prepared using azolectin in decane containing 1% chloroform (Sigma) across a 250 μM hole in a polystyrene cuvette. Azolectin was partially purified by precipitation with cold acetone from a chloroform solution. The inside of the cuvette constituted the *trans* compartment. The standard experimental medium was 100 mM $CaCl_2$ (or Ca-gluconate), 10 mM HEPES/pH 7.2. The lipid membrane was built under symmetric ionic conditions and both chambers contained 3 ml of solution. The contents of both chambers were stirred by magnetic bars when desired. Connections to the electrodes were provided by agar bridges. Purified proteins were added to the *cis* side. Control experiments with empty membrane or with detergents used for the purification showed no activity. All voltages reported are those of the *cis* chamber, zero being assigned by convention to the *trans* (grounded) side. Currents are considered as positive when carried by cations flowing from the *cis* to the *trans* compartment. A BC-525C unit and headstage were used to control parameters and amplify signals. Output was recorded with a 10 kHz bandwidth on videotape using a Medical Systems PCM-2 interface. Data were acquired at 2 kHz, filtered at 500 Hz and analysed offline using the pClamp program set (Axon Instruments). Conductance was determined by averaging the measured amplitudes of single channel events ($n \geq 50$) at various applied voltages.

**Statistical analysis of data.** Statistical data are presented as mean ± s.e.m., unless otherwise specified. Significance was calculated by Student's *t*-test, and correlation analysis was performed with the SigmaPlot 11.0 software (Systat Software Inc.).


27. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305,** 567–580 (2001).
28. Schoenmakers, T. J. M., Visser, G. J., Flik, G. & Theuvenet, A. P. R. CHELATOR: an improved method for computing metal ion concentrations in physiological solutions. *Biocomputing* **12,** 870–879 (1992).
29. Bolte, S. & Cordelieres, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224,** 213–232 (2006).
30. Costes, S. V. *et al.* Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophys. J.* **86,** 3993–4003 (2004).
31. Szabò, I., Soddemann, M., Leanza, L., Zoratti, M. & Gulbins, E. Single-point mutations of a lysine residue change function of Bax and Bcl-$x_L$ expressed in Bax- and Bak-less mouse embryonic fibroblasts: novel insights into the molecular mechanisms of Bax-induced apoptosis. *Cell Death Differ.* **18,** 427–438 (2011).
32. Teardo, E. *et al.* Characterization of a plant glutamate receptor activity. *Cell. Physiol. Biochem.* **26,** 253–262 (2010).


©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10234

# Integrative genomics identifies MCU as an essential component of the mitochondrial calcium uniporter

Joshua M. Baughman[1,2]*, Fabiana Perocchi[1,2]*, Hany S. Girgis[1,2], Molly Plovanich[1,2], Casey A. Belcher-Timme[1,2], Yasemin Sancak[1,2], X. Robert Bao[1,2], Laura Strittmatter[1,2], Olga Goldberger[1,2], Roman L. Bogorad[3], Victor Koteliansky[4] & Vamsi K. Mootha[1,2]

**Mitochondria from diverse organisms are capable of transporting large amounts of $Ca^{2+}$ via a ruthenium-red-sensitive, membrane-potential-dependent mechanism called the uniporter[1–4]. Although the uniporter's biophysical properties have been studied extensively, its molecular composition remains elusive. We recently used comparative proteomics to identify MICU1 (also known as CBARA1), an EF-hand-containing protein that serves as a putative regulator of the uniporter[5]. Here, we use whole-genome phylogenetic profiling, genome-wide RNA co-expression analysis and organelle-wide protein coexpression analysis to predict proteins functionally related to *MICU1*. All three methods converge on a novel predicted transmembrane protein, CCDC109A, that we now call 'mitochondrial calcium uniporter' (MCU). MCU forms oligomers in the mitochondrial inner membrane, physically interacts with MICU1, and resides within a large molecular weight complex. Silencing *MCU* in cultured cells or *in vivo* in mouse liver severely abrogates mitochondrial $Ca^{2+}$ uptake, whereas mitochondrial respiration and membrane potential remain fully intact. MCU has two predicted transmembrane helices, which are separated by a highly conserved linker facing the intermembrane space. Acidic residues in this linker are required for its full activity. However, an S259A point mutation retains function but confers resistance to Ru360, the most potent inhibitor of the uniporter. Our genomic, physiological, biochemical and pharmacological data firmly establish MCU as an essential component of the mitochondrial $Ca^{2+}$ uniporter.**

To predict proteins that are functionally related to MICU1 (ref. 5) and essential for mitochondrial calcium ($Ca^{2+}$) uptake, we performed three systematic computational analyses. First, we ranked all ~20,000 mammalian genes on the basis of the similarity of their phylogenetic profile to *MICU1*, where the phylogenetic profile of a gene is defined as the binary vector of presence or absence of its homologues across 500 evolutionarily diverse organisms[6,7]. Second, we scored all ~20,000 mammalian genes for their expression similarity to *MICU1* across 81 mouse cell types and tissues using a genome-wide RNA expression atlas[8]. Third, we scored protein expression similarity between MICU1 and all ~1,100 mitochondrial proteins, based on their pattern of peptide abundance across 14 different mouse tissues[6]. All three computational methods (Fig. 1a–c) spotlight an unstudied protein (previously called CCDC109A, accession number NM_138357.1) that we now call 'mitochondrial calcium uniporter' (MCU). MCU, which has two predicted transmembrane domains, was first discovered in our proteomic analysis as a mitochondrial protein detected in 12 different mouse tissues[6]. *MCU*, which bears no sequence similarity to *MICU1*, is tied as the eighth closest phylogenetic neighbour genome-wide (Fig. 1a), being either co-present or co-absent with *MICU1* in 495 of 500 organisms evaluated (Hamming distance = 5). *MCU* is also the second highest scoring gene in the genome-wide mRNA co-expression analysis (Fig. 1b) and is the top scoring protein amongst all ~1,100 mitochondrial proteins for protein coexpression with MICU1 across 14 mouse tissues (Fig. 1c).

Our prediction of a functional relationship between MICU1 and MCU was further corroborated by evidence of a physical interaction. By transfecting green-fluorescent-protein-tagged MCU (MCU–GFP) into cells stably expressing V5-tagged MICU1 (MICU1–V5) and vice versa, we were able to recover both GFP-tagged proteins following

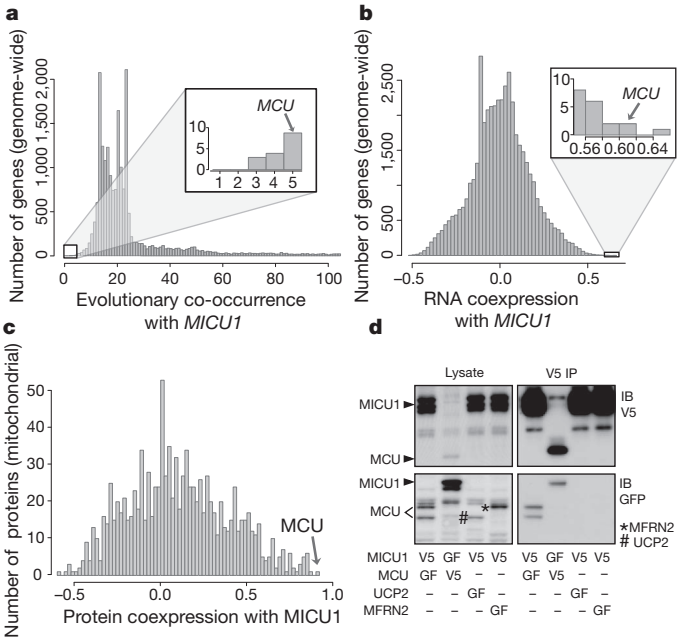



**Figure 1 | Integrative genomics predicts MCU to be functionally related to MICU1** **a–c**, Phylogenetic profile neighbours (**a**), RNA co-expression neighbours (**b**) and protein co-expression neighbours (**c**) of MICU1. Hamming distances between phylogenetic profiles were computed genome-wide for all 20,000 mammalian genes across 500 fully sequenced organisms. Genes co-expressed with *MICU1* were computed genome-wide by Pearson correlation using a mouse atlas of 81 tissues. Protein expression correlation with MICU1 was analysed for all mitochondrial proteins across 14 mouse tissues. **d**, Coimmunoprecipitation of MICU1 and MCU. HEK-293 cells stably expressing MICU1–V5 or MCU–V5 were transfected with MICU1–GFP, MCU–GFP, MFRN2–GFP or UCP2–GFP. Cell lysates were incubated with anti-V5 antibody, immunoprecipitates were resolved on SDS–PAGE, and input lysates and immunoprecipitates were blotted with anti-V5 (top) or anti-GFP (bottom) antibodies. Data are representative of three independent experiments. IP, immunoprecipitation; GF, green fluorescent protein; IB, immunoblot.

[1]Departments of Systems Biology and Medicine, Harvard Medical School and Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [2]Broad Institute, Cambridge, Massachusetts 02142, USA. [3]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [4]Alnylam Pharmaceuticals, Inc., Cambridge, Massachusetts 02142, USA.

*These authors contributed equally to this work.



©2011 Macmillan Publishers Limited. All rights reserved

immunoprecipitation with an anti-V5 antibody (Fig. 1d). The interaction was specific as MICU1–V5 was incapable of pulling down two different GFP-tagged inner membrane proteins (UCP2 and MFRN2, also known as SLC25A28). Similar results were also obtained by immunoprecipitating Flag-tagged MCU and probing for endogenous MICU1 (Supplementary Fig. 1).

Collectively, our three complementary genomic analyses combined with our biochemical data (Fig. 1) predict that MCU is functionally related to MICU1 and that it, too, may participate in mitochondrial $Ca^{2+}$ uptake.

We evaluated the impact of silencing *MCU* on mitochondrial $Ca^{2+}$ uptake in intact and permeabilized cells using RNA interference (RNAi). Silencing *MCU* in a HeLa cell line expressing a mitochondria-targeted aequorin (mt-Aeq) reporter[9] attenuates mitochondrial $Ca^{2+}$ uptake (Fig. 2a) proportionate to the strength of knockdown (Fig. 2b). The RNAi-induced phenotype is not off-target because coexpression of a full-length *MCU* cDNA together with a short hairpin RNA (shRNA) that targets the *MCU* 3′ untranslated region (3′UTR) fully rescues $Ca^{2+}$ uptake (Fig. 2a). Moreover, the RNAi effect is not a trivial consequence of interrupting upstream signalling, because histamine mobilization of cytosolic $Ca^{2+}$ remains intact (Supplementary Fig. 2a), and because we obtain similar results when measuring clearance of exogenously added $Ca^{2+}$ by mitochondria in permeabilized HEK-293 (Fig. 2c) and HeLa cells (Supplementary Fig. 3). In HeLa cells, basal and uncoupled respiration were intact (Supplementary Fig. 2b), mitochondrial membrane potential ($\psi_m$) was not depolarized (Supplementary Fig. 2c), and mitochondrial morphology remained grossly unchanged (data not shown) after silencing of *MCU*. Basal cytosolic $Ca^{2+}$ levels were 53.9 nM ± 16.3 ($n = 30$) in *MCU* knockdown HeLa cells and 70.9 nM ± 11.7 ($n = 14$) in control sh-LACZ cells. Although mitochondrial $Ca^{2+}$ buffering is known to shape cytosolic $Ca^{2+}$ transients in many cell types, its inhibition through silencing of *MCU* did not show a significant impact on cytosolic $Ca^{2+}$ clearance after histamine stimulation (Supplementary Fig. 2a). However, this result could reflect an incomplete silencing of *MCU*-mediated mitochondrial $Ca^{2+}$ uptake in HeLa cells (Fig. 2a, b). Mitochondrial $Ca^{2+}$ uptake has also previously been shown to stimulate ATP production by allosterically activating three tricarboxylic acid (TCA) cycle dehydrogenases[10]. Using a previously reported protocol[11], we found that permeabilized sh-MCU cells exhibit attenuated NAD(P)H elevation in response to exogenously added $Ca^{2+}$ (Supplementary Fig. 2d), demonstrating that silencing *MCU*, like *MICU1* (ref. 5), attenuates $Ca^{2+}$ activation of the TCA cycle.

To complement these cell-based studies, we analysed mitochondrial calcium uptake in mouse liver mitochondria upon *in vivo* silencing of *MCU*. A key advantage of this experimental system is that $Ca^{2+}$ uptake phenotypes are directly attributable to mitochondria, and additional bioenergetic parameters, notably respiratory flux and $\psi_m$, can be carefully measured under classic respiratory state transitions[12]. As respiratory state transitions are dependent on TCA cycle metabolism, electron transport, $H^+$ pumping, ATP/ADP exchange, and transport of phosphate, such studies allow us to evaluate the specificity of MCU's role in mitochondrial $Ca^{2+}$ transport. Using previously described siRNA design and delivery technology[13], we screened 46 distinct siRNA duplexes. We selected a duplex characterized by a half-maximal inhibitory concentration of 69 pM that achieved >90% knockdown at a concentration of 5 nM in cultured mouse liver cells (Fig. 2d). We performed a large-scale synthesis of this siRNA duplex and encapsulated it into a lipid-based formulation optimized for liver-specific delivery[14,15]. As a negative control we used a siRNA duplex specific for the firefly luciferase gene (si-LUC). We performed weekly tail vein injections of the siRNA duplexes over a 3 week period and achieved 84 ± 2% ($n = 3$) *in vivo* mRNA knockdown of *MCU* in the liver, which was corroborated by immunoblot analysis (Fig. 2d). Both groups of mice showed normal weight gain (5.6 ± 1.3% of body weight for si-LUC and 5.0 ± 3.0% of body weight for si-MCU, $n = 3$) and the gross appearance of the livers did not differ. Mitochondria isolated from

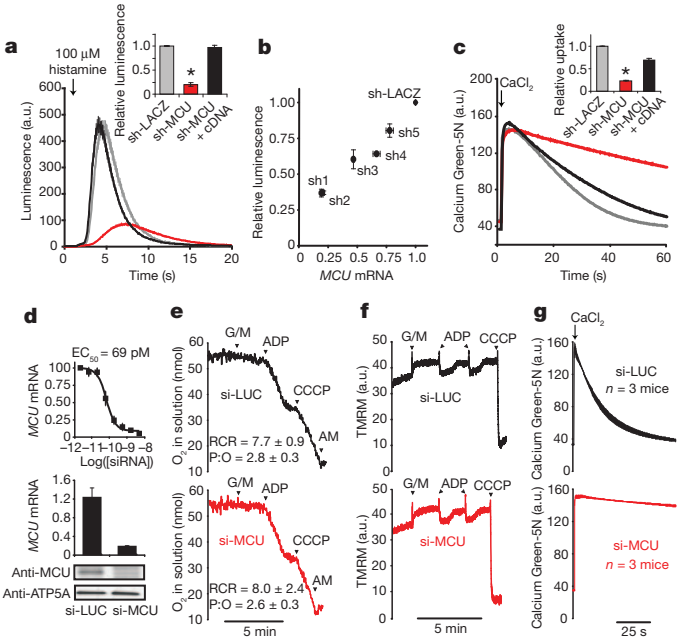



**Figure 2 | MCU is required for mitochondrial $Ca^{2+}$ uptake in cultured cells and in purified mouse liver mitochondria.** **a**, Representative luminescence measurements of a mitochondrial aequorin $Ca^{2+}$ reporter after histamine stimulation in HeLa cells expressing sh-LACZ, sh-MCU, or a combination of sh-MCU and an RNAi-resistant cDNA for *MCU* (mean ± s.e.m., $n = 10$ traces). Inset shows statistical analysis of the maximal luminescence (mean ± s.d., $n = 10$ traces, $^*P < 0.001$). a.u., arbitrary units. **b**, Relationship between *MCU* mRNA expression and histamine-induced mitochondrial $Ca^{2+}$ uptake (maximal aequorin luminescence) recorded from five independent shRNAs targeting *MCU* and normalized to sh-LACZ (mean ± s.d., $n = 3$). **c**, Representative traces of $Ca^{2+}$ uptake in digitonin-permeabilized sh-MCU HEK-293 cells or sh-LACZ control cells after addition of 50 μM final concentration of $CaCl_2$. Inset reports linear fits of uptake kinetics between 15 and 20 s, normalized to sh-LACZ (mean ± s.d., $n = 3$, $^*P < 0.001$). $Ca^{2+}$ was measured with Calcium Green-5N. **d**, *In vitro* dose–response of a selected siRNA duplex targeting mouse *MCU*. Relative expression of *MCU* mRNA in livers following weekly injections of si-MCU or si-LUC for 3 weeks, normalized to expression in PBS-treated mice. Immunoblot analysis of liver mitochondria isolated from mice treated with si-MCU or control si-LUC using antibodies against MCU and ATP5A1 as a loading control. **e**, Oxygen consumption measurements of isolated mitochondria in a well-stirred cuvette. Glutamate and malate (G/M), ADP, uncoupler (carbonyl cyanide *m*-chlorophenyl hydrazone, CCCP), antimycin (AM) were added at indicated time points. Respiratory control ratio (RCR) and ADP:O ratio (P:O) were computed from three separate mice for each group. **f**, Mitochondrial membrane potential ($\psi_m$) measured by tetramethyl rhodamine methyl ester (TMRM) in isolated liver mitochondria. **g**, $Ca^{2+}$ uptake kinetics in energized liver mitochondria following the addition of 50 μM final $CaCl_2$. Extra-mitochondrial $Ca^{2+}$ was measured with Calcium Green-5N (mean ± s.e.m., $n = 3$ mice). Traces depicted in **e** and **f** are representative of measurements made from three independent mouse experiments performed on separate days.

these livers were physiologically intact, capable of undergoing robust respiratory state transitions (Fig. 2e, f). The respiratory control ratio (RCR) was comparable in si-LUC (7.7 ± 0.9, $n = 3$) and si-MCU animals (8.0 ± 2.4, $n = 3$). The ADP:O ratios were 2.77 ± 0.3 ($n = 3$) and 2.61 ± 0.3 ($n = 3$) in si-LUC and si-MCU mitochondria, respectively, indicating intact respiratory chain physiology. However, mitochondria from si-MCU mice show a profound and near complete loss in $Ca^{2+}$ uptake in response to extramitochondrial pulses of $Ca^{2+}$ (Fig. 2g and Supplementary Fig. 4).

Collectively, our physiology studies, comprising three different experimental systems (intact cells, permeabilized cells and mitochondria purified from mouse organs), three different cell types (HeLa, HEK-293 and mouse liver), and two different silencing modalities (*in vitro* lentiviral shRNA and *in vivo* siRNA), firmly establish that



©2011 Macmillan Publishers Limited. All rights reserved

*MCU* is essential for high capacity $Ca^{2+}$ transport into mitochondria, and that the phenotype is not a secondary consequence of alterations in cytosolic $Ca^{2+}$ signalling or of impaired ion or metabolite transport across the mitochondrial inner membrane. To our knowledge, this is the first time that mitochondrial $Ca^{2+}$ uptake has been silenced *in vivo* in an animal.

We next established MCU's precise sub-organellar localization and topology. MCU was first identified as a mitochondrial protein in our previous proteomic survey[6], where it was detected as a low abundance protein in 12 out of 14 mouse tissues, being missed only in heart and kidney. Confirmatory immunoblot analysis suggests it is also present in the heart and the kidney (data not shown), indicating it is universally expressed. As expected, a carboxy terminus GFP-tagged version of MCU localizes exclusively to mitochondria in HeLa cells (Fig. 3a) as well as in HEK-293 cells (data not shown). To biochemically validate MCU's localization, we fractionated HeLa cell lysate and observed significant enrichment of the approximately 35 kDa native MCU in the mitochondrial fraction (Fig. 3b). The observed molecular weight of MCU is lower than the predicted 40 kDa and is consistent with a predicted cleavable amino-terminal mitochondrial targeting sequence[16]. Multiple algorithms predict that MCU contains two transmembrane domains[17,18]. To experimentally confirm mitochondrial inner-membrane integration and to discriminate between transmembrane and peripheral membrane configurations, we performed alkaline carbonate extraction of proteins from intact mitochondria isolated from HEK-293 cells expressing MCU-V5. We find that MCU remains in the insoluble fraction, proving that it is a bona fide transmembrane protein (Fig. 3c). To establish the topology of MCU across the inner membrane, we disrupted the mitochondrial outer membrane by hypotonic swelling and treated the resulting mitoplasts with proteinase K. Proteins such as OXA1 that are exposed to the mitochondrial intermembrane space are digested by proteinase K, whereas COII (an integral inner membrane protein, also known as MT-CO2) and MCU are completely protected (Fig. 3d), despite being substrates for proteinase K (Supplementary Fig. 5). Our data therefore indicate that MCU is localized to the inner membrane with both its N- and C-termini facing the matrix space.

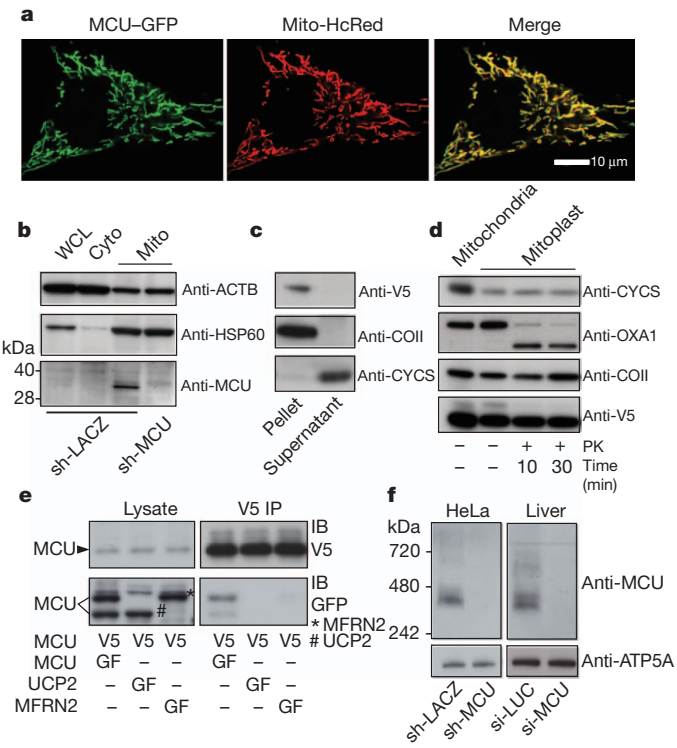



**Figure 3 | MCU is oligomeric and resides in the mitochondrial inner membrane as a larger complex.** **a**, Confocal imaging of MCU–GFP co-expressed with mitochondria-targeted HcRed (Mito-HcRed) in HeLa cells. **b**, Immunoblot analysis of HeLa whole-cell lysate (WCL), cytosol (Cyto) or crude mitochondrial fractions (Mito), using antibodies against MCU, HSP60 (matrix protein, also known as HSPD1), or ACTB (cytosol). **c**, Immunoblot analysis of soluble (supernatant) and insoluble (pellet) fractions following alkaline carbonate extraction of mitochondrial fractions from HEK-293 cells expressing MCU–V5. Immunoblot analysis was performed using antibodies against V5, COII (integral inner membrane protein) and CYCS (soluble intermembrane space protein). **d**, Immunoblot analysis after proteinase K (PK) treatment of MCU–V5-expressing HEK-293 mitoplasts for indicated times. **e**, Anti-V5 immunoprecipitations performed as in Fig. 1d using lysates from HEK-293 cells stably expressing MCU–V5 and transiently transfected with MCU–GFP, UCP2–GFP, or MFRN2–GFP. **f**, Blue native PAGE analysis of mitochondrial fractions from HeLa cells (stably expressing sh-LACZ or sh-MCU, left panel) or from livers of mice (si-LUC or si-MCU, right panel) and immunoblotted for MCU. ATP5A1 is used as a loading control.

If MCU is an integral component of the uniporter, we might expect that it oligomerizes and operates within a larger complex. To test for MCU oligomerization, we performed pull-down experiments and found that V5-tagged MCU is capable of immunoprecipitating GFP-tagged MCU (Fig. 3e), indicating that MCU can form multimers. To control for nonspecific interactions we used two inner mitochondrial membrane spanning transporters as controls, neither of which immunoprecipitated with MCU–V5 (Fig. 3e). We performed blue native gel separation of digitonin-solubilized mitochondria purified from HeLa cells and found that MCU migrates at an apparent molecular weight of ~450 kDa. This larger complex disappears following silencing with sh-MCU (Fig. 3f, left panel). We obtained similar results using mitochondria from livers of si-LUC and si-MCU mice (Fig. 3f, right panel). Collectively, these studies indicate that MCU oligomerizes in the mitochondrial inner membrane as part of a larger molecular weight complex.

Our topology analysis (Fig. 3c, d) combined with computational predictions of membrane spanning domains indicate that MCU's two transmembrane helices, TM1 and TM2, are linked by a short stretch of amino acids that face the intermembrane space and contain what we now term a 'DIME' motif (Fig. 4a). We sought to determine whether conserved amino acids near and within the DIME motif are required for MCU-mediated $Ca^{2+}$ uptake. We created a series of four alanine mutants at conserved residues (E257A, S259A, D261A, E264A) and expressed them on an sh-MCU background to evaluate their ability to rescue $Ca^{2+}$ transport (Fig. 4b). One mutant, S259A, was capable of providing significant rescue of the sh-MCU phenotype (Fig. 4b). The remaining three mutants (E257A, D261A, E264A), however, failed to restore mitochondrial calcium uptake despite comparable expression to the S259A mutant (Supplementary Fig. 6a), signifying that these highly conserved acidic residues are critical for $Ca^{2+}$ transport.

Ruthenium red and a related compound, Ru360, are the most potent inhibitors of the mitochondrial $Ca^{2+}$ uniporter, yet their target remains unknown[19,20]. These cell-impermeant inhibitors have been useful pharmacological tools in studies of isolated mitochondria[21,22] and are believed to act on the outer face of the inner membrane. We next sought to determine whether the functional S259A mutant might influence Ru360 sensitivity. Ru360 strongly abolishes mitochondrial $Ca^{2+}$ uptake in permeabilized HEK-293 cells (Fig. 4c). Transient overexpression of *MCU* leads to a mild resistance to Ru360 compared to control (Fig. 4c). However, transient expression of the S259A mutant confers marked resistance to Ru360 in both HEK-293 and HeLa cells (Fig. 4c, Supplementary Fig. 6b, c, d, e). These experiments clearly indicate a functional role for the linker region in both activity and Ru360 sensitivity of MCU.

Our genomic, physiological, biochemical and pharmacological data firmly establish that MCU is an oligomeric, mitochondrial inner membrane protein that is essential for calcium uniporter activity. In our current studies, overexpression of *MCU* alone did not give rise to a



©2011 Macmillan Publishers Limited. All rights reserved

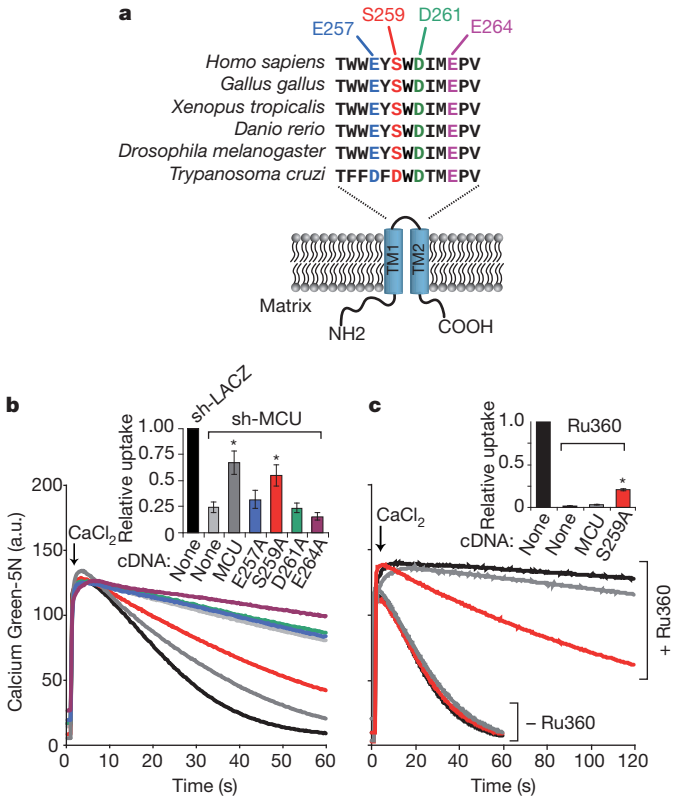



**Figure 4 | Impact of point mutations on MCU activity and its sensitivity to Ru360.** **a**, Schematic of MCU topology across the mitochondrial inner membrane and a multiple sequence alignment of the linker sequence containing a DIME motif. TM1 and TM2 are two transmembrane domains. **b**, $Ca^{2+}$ uptake in permeabilized sh-MCU HEK-293 cells transiently expressing MCU mutants. Inset reports linear fits of uptake kinetics between 15 and 25 s, normalized to sh-LACZ (mean ± s.d., $n = 3$, $^*P < 0.01$) **c**, $Ca^{2+}$ uptake in HEK-293 cells transiently expressing wild-type MCU or the S259A mutant, in the presence or absence of 0.5 μM Ru360. Inset reports linear fits of uptake kinetics between 30 and 60 s for Ru360-treated cells, and between 15 and 25 s for untreated cells. Uptake rates are normalized to untreated HEK-293 cells (mean ± s.d., $n = 3$, $^*P < 0.01$).

marked gain of $Ca^{2+}$ uptake in HeLa cells (Fig. 2a), indicating that additional components or chaperones may be limiting *in vivo*. Although its precise mechanism remains to be elucidated, MCU fulfils multiple criteria for being an integral component of the uniporter, where it could serve as the pore-forming subunit or as a closely associated partner protein. First, loss of *MCU* leads to a profound defect in mitochondrial $Ca^{2+}$ uptake in three distinct systems: intact cells, permeabilized cells and mouse liver (Fig. 2). Second, it has two bona fide transmembrane domains and oligomerizes to form a larger molecular weight complex in the mitochondrial inner membrane (Fig. 3e, f). Third, mutations at evolutionarily conserved acidic residues attenuate its activity (Fig. 4b). Fourth, a point mutation in MCU confers resistance against Ru360, the most potent inhibitor of the uniporter (Fig. 4c and Supplementary Fig. 6). The current study further supports our previously reported phylogenomic signature of the mitochondrial $Ca^{2+}$ uniporter[5], and together these studies identify MICU1 and MCU as an evolutionarily conserved duo that are critical for mitochondrial $Ca^{2+}$ uptake. Whereas MICU1 may serve as a regulator, MCU is likely to represent a core component of the uniporter. Future studies will be necessary to establish the exact mechanism by which these proteins function together to orchestrate the mitochondrial $Ca^{2+}$ cycle.

## METHODS SUMMARY

Candidate human genes required for mitochondrial $Ca^{2+}$ uptake were prioritized on the basis of the co-occurrence of their homologues with *MICU1* across 500 evolutionarily diverse organisms[6,7] and their RNA and protein co-expression across mouse cell types and tissues[6,8]. Stable knockdown of *MCU* in HeLa and HEK-293 cells was achieved using lentiviral shRNA-expressing constructs available from the Broad Institute RNAi Consortium[23]. *In vivo* silencing of *MCU* in mouse liver was achieved via weekly tail vein injections of selected siRNAs in lipid-based formulations[13–15,24]. cDNA rescue studies in *MCU* knockdown cells were carried out by overexpression of wild-type or mutant versions of full-length MCU cDNAs together with a shRNA that targets the *MCU* 3′UTR. Agonist-induced rises in mitochondrial $Ca^{2+}$ in intact HeLa cells were measured by luminescence of a mitochondria-targeted aequorin reporter[9]. Extra-mitochondrial calcium in permeabilized HeLa cells, HEK-293 and isolated liver mitochondria was measured using Calcium Green-5N (ref. 5). Single-cell studies of cytosolic calcium were performed as previously described[25]. Mitochondrial respiration, membrane potential and NAD(P)H were measured via established protocols[11,26]. Crude mitochondria and mitoplasts were prepared from cultured HEK-293 cells expressing a C terminus V5-tagged version of MCU[5]. Protein topology was assessed by alkaline carbonate extraction from crude mitochondria and proteinase K digestion of mitoplasts[27]. Protein interaction studies were performed by immunoprecipitation with anti-V5 and anti-Flag antibodies. Native gel electrophoresis was performed as previously described[28]. Unless otherwise indicated, data are summarized as mean ± standard deviation (s.d.), and *P*-values were computed from *t*-tests.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 9 November 2010; accepted 2 June 2011.**
**Published online 19 June 2011.**

1. DeLuca, H. F. & Engstrom, G. W. Calcium uptake by rat kidney mitochondria. *Proc. Natl Acad. Sci. USA* **47,** 1744–1750 (1961).
2. Vasington, F. D. & Murphy, J. V. $Ca^{++}$ ion uptake by rat kidney mitochondria and its dependence on respiration and phosphorylation. *J. Biol. Chem.* **237,** 2670–2677 (1962).
3. Carafoli, E. & Lehninger, A. L. A survey of the interaction of calcium ions with mitochondria from different tissues and species. *Biochem. J.* **122,** 681–690 (1971).
4. Gunter, K. & Gunter, T. E. Transport of calcium by mitochondria. *J. Bioenerg. Biomembr.* **26,** 471–485 (1994).
5. Perocchi, F. *et al. MICU1* encodes a mitochondrial EF hand protein required for $Ca^{2+}$ uptake. *Nature* **467,** 291–296 (2010).
6. Pagliarini, D. J. *et al.* A mitochondrial protein compendium elucidates complex I disease biology. *Cell* **134,** 112–123 (2008).
7. Pellegrini, M., Marcotte, E. M., Thompson, M. J., Eisenberg, D. & Yeates, T. O. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc. Natl Acad. Sci. USA* **96,** 4285–4288 (1999).
8. Lattin, J. E. *et al.* Expression analysis of G Protein-Coupled Receptors in mouse macrophages. *Immunome Res.* **4,** 5 (2008).
9. Rizzuto, R., Simpson, A. W., Brini, M. & Pozzan, T. Rapid changes of mitochondrial $Ca^{2+}$ revealed by specifically targeted recombinant aequorin. *Nature* **358,** 325–327 (1992).
10. Denton, R. M. & McCormack, J. G. The role of calcium in the regulation of mitochondrial metabolism. *Biochem. Soc. Trans.* **8,** 266–268 (1980).
11. Territo, P. R., Mootha, V. K., French, S. A. & Balaban, R. S. $Ca^{2+}$ activation of heart mitochondrial oxidative phosphorylation: role of the $F_0/F_1$-ATPase. *Am. J. Physiol. Cell Physiol.* **278,** C423–C435 (2000).
12. Chance, B. & Williams, G. R. Respiratory enzymes in oxidative phosphorylation. I. Kinetics of oxygen utilization. *J. Biol. Chem.* **217,** 383–393 (1955).
13. Musunuru, K. *et al.* From noncoding variant to phenotype via *SORT1* at the 1p13 cholesterol locus. *Nature* **466,** 714–719 (2010).
14. Akinc, A. *et al.* A combinatorial library of lipid-like materials for delivery of RNAi therapeutics. *Nature Biotechnol.* **26,** 561–569 (2008).
15. Akinc, A. *et al.* Targeted delivery of RNAi therapeutics with endogenous and exogenous ligand-based mechanisms. *Mol. Ther.* **18,** 1357–1364 (2010).
16. Claros, M. G. & Vincens, P. Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem.* **241,** 779–786 (1996).
17. Bernsel, A., Viklund, H., Hennerdal, A. & Elofsson, A. TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Res.* **37,** W465–W468 (2009).
18. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305,** 567–580 (2001).
19. Moore, C. L. Specific inhibition of mitochondrial $Ca^{++}$ transport by ruthenium red. *Biochem. Biophys. Res. Commun.* **42,** 298–305 (1971).
20. Ying, W. L., Emerson, J., Clarke, M. J. & Sanadi, D. R. Inhibition of mitochondrial calcium ion transport by an oxo-bridged dinuclear ruthenium ammine complex. *Biochemistry* **30,** 4949–4952 (1991).
21. Bernardi, P. Mitochondrial transport of cations: channels, exchangers, and permeability transition. *Physiol. Rev.* **79,** 1127–1155 (1999).
22. Kirichok, Y., Krapivinsky, G. & Clapham, D. E. The mitochondrial calcium uniporter is a highly selective ion channel. *Nature* **427,** 360–364 (2004).



©2011 Macmillan Publishers Limited. All rights reserved

23. Moffat, J. *et al.* A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124,** 1283–1298 (2006).
24. Semple, S. C. *et al.* Rational design of cationic lipids for siRNA delivery. *Nature Biotechnol.* **28,** 172–176 (2010).
25. Rebres, R. A. *et al.* Synergistic $Ca^{2+}$ responses by $G\alpha_i$- and $G\alpha_q$-coupled G-protein-coupled receptors require a single PLCβ isoform that is sensitive to both $G\beta\gamma$ and $G\alpha_q$. *J. Biol. Chem.* **286,** 942–951 (2011).
26. Gohil, V. M. *et al.* Nutrient-sensitized screening for drugs that shift energy metabolism from mitochondrial respiration to glycolysis. *Nature Biotechnol.* **28,** 249–255 (2010).
27. Ryan, M. T., Voos, W. & Pfanner, N. Assaying protein import into mitochondria. *Methods Cell Biol.* **65,** 189–215 (2001).
28. Schagger, H. & von Jagow, G. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **199,** 223–231 (1991).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank R. Nilsson, J. Engreitz and S. Calvo for bioinformatics assistance; D. Root and S. Silver for assistance in lentiviral RNAi; B. R. Bettencourt, K. Charisse, S. Kuchimanchi and L. Speciner for siRNA design, synthesis and formulation; M. Blower, J. Avruch and R. Ward for advice; and members of the Mootha laboratory for valuable feedback. J.M.B. and L.S. were supported by graduate student fellowships from the National Science Foundation. This work was supported by grants from the National Institutes of Health (GM0077465, DK080261) awarded to V.K.M.

**Author Contributions** J.M.B., F.P. and V.K.M. conceived of the project and its design. J.M.B., F.P., H.S.G., M.P., O.G., L.S., C.A.B.-T., X.R.B., Y.S. and R.L.B. performed experiments and data analysis. V.K. aided in experimental design. V.K.M., J.M.B., F.P. and M.P. wrote the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to V.K.M. (vamsi@hms.harvard.edu).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Phylogenetic profiling and co-expression analyses.** We performed genome-wide phylogenetic profiling using a previously reported phylogenetic matrix across 500 sequenced genomes using the Hamming distance as our similarity metric[6]. For genome-wide mRNA co-expression analysis, we used a publicly available genome-wide tissue gene expression atlas (NCBI GSE10246)[8]. We used the published gene expression intensities normalized to the median for each chip, and then computed the Pearson correlation to *MICU1* expression. For protein co-expression analysis, we used our MitoCarta protein atlas of mitochondrial protein expression across 14 tissues[6]. We summarized the abundance of a protein as the number of detected peptides per protein per tissue, and computed the Pearson correlation to MICU1 expression.

**Cell culture.** HEK-293 cells, Hepa-1c1c7 and HeLa cells were received from the ATCC. HeLa cells expressing a mitochondrial matrix targeted aequorin (mt-Aeq) were purchased from Aequotech (AT-002-H). All cells were grown in Dulbecco's modified Eagle medium (DMEM) high glucose medium (Invitrogen) with 10% FBS (Sigma) at 37 °C and 5% $CO_2$. mt-Aeq cells were maintained in 100 μg ml$^{-1}$ geneticin (Gibco).

**Plasmids and reagents.** Lentiviral vectors for expressing shRNA (pLKO.1) or V5-tagged cDNAs (pLEX983) were obtained from the Broad Institute's RNAi Consortium (TRC). To silence *MCU*, we used the following hairpins from the TRC: sh1 (TRCN0000137529, 5′-CCAGCAACTATACACCACACT-3′), sh2 (TRCN0000133861, 5′-GCAAGGAGTTTCTTTCTCTTT-3′), sh3 (TRCN0000133984, 5′-CAATCAACTCAAGGATGCAAT-3′), sh4 (TRCN0000135430, 5′-GCCATG GCAATGTATGCATAT-3′), sh5 (TRCN0000138929, 5′-GATCGCTTCCTGG CAGAATTT-3′). For transfections, cDNAs were inserted into pcDNA vectors with C-terminal V5/His6× or GFP tags (Invitrogen pDEST40 and pDEST47, respectively) using Gateway LR (Invitrogen). A full-length *MCU* human cDNA (NM_138357.1) without a stop codon was cloned into the pDONR223 Gateway vector and subsequently cloned into pLEX983, pDEST40 and pDEST47 vectors. MCU mutants E257A, S259A, D261A and E264A mutations were created using Stratagene QuikChange mutagenesis. We used TaqMan assays (ABI) to quantify the *MCU* transcript (*CCDC109A*).

**Lentivirus production and infection.** Procedures and reagents for virus production and infection were adapted from the Broad RNAi Consortium protocols and were described previously[23]. Selection was begun 24 h post-infection with 2 μg ml$^{-1}$ puromycin or 5 μg ml$^{-1}$ blasticidin.

**Rescue of sh-*MCU* and overexpression of MCU mutants.** HeLa cells stably overexpressing shRNAs targeting either *LACZ* or the 3′UTR of *MCU* (TRCN0000133861) were transduced with either pLEX983-MCU wild-type or mutant versions and selected with 5 μg ml$^{-1}$ blasticidin for at least 1 week before assaying. Wild-type HEK-293 or HEK-293 cells stably overexpressing sh-MCU (TRCN0000133861) were transfected with wild-type or mutant MCU cDNAs using Fugene6 reagent from Roche. To ensure high transfection efficiency HEK-293 were transfected twice 3 days apart and $Ca^{2+}$ uptake was assayed 7 days after the first transfection.

**Synthesis and *in vitro* screening of siRNAs targeting *MCU*.** siRNAs with the lowest predicted off-target potentials (46) and 100% homology with mouse sequence NM_001033259.3 were selected for synthesis and screening. Single-strand RNAs were produced at Alnylam Pharmaceuticals as previously described[13]. MCU siRNA duplex with lowest $EC_{50}$ and lowest predicted off-target potential was selected. The siRNA sense strand is 5′-GAcuGAGAGA cccAuuAcAdTsdT-3′, antisense is 5′-UGuAAUGGGUCUCUcAGUCdTsdT-3′; control luciferase siRNA sense is 5′-cuuAcGcuGAGuAcuucGAdTsdT-3′, and antisense is 5′-UCGAAGuACUcAGCGuAAGdTsdT-3′. Chemical modifications were introduced to stabilize siRNA *in vivo*, reduce off-target potential of sense strand and minimize immune response (2′-OMe modified nucleotides are in lower case, and phosphorothioate linkages are indicated by 's'). Hepa-1c1c7 cells (seeded 15,000 cells per well in 96-well plates) were transfected with siRNA using the Lipofectamine RNAiMAX reagent according to the manufacturer's protocols. *MCU* mRNA levels were quantified in cell lysates 18–24 h post-transfection using a branched-DNA assay (QuantiGene Reagent System, Panomics), according to the manufacturer's protocols. *MCU* mRNA levels were normalized to *GAPDH* mRNA.

***In vivo* silencing of *MCU*.** All procedures used in animal studies were approved by the Institutional Animal Care and Use Committee. C57BL/6 mice (Charles River laboratories) received either PBS or siRNA in lipid-based formulations (as previously described[14,15,24]) via tail vein injection weekly. After overnight fasting the animals were euthanized by $CO_2$ inhalation and liver tissues were collected into ice-cold PBS for mitochondria isolation. A piece of liver tissue was snap-frozen in liquid nitrogen for measurements of mRNA and protein levels.

**Luminescence-based measurement of mitochondrial $Ca^{2+}$.** Assays were performed using either mt-Aeq HeLa cells that stably express mitochondrial aequorin (Fig. 2b) or wild-type HeLa cells that were transduced with an equal titre of lentivirus to express mitochondrial-targeted aequorin[9] 3 days before assaying (Fig. 2a). 12–18 h before assaying, 50,000 cells were seeded per well in a 96-well plate. Light emission after histamine stimulation was measured as reported previously[5]. After histamine stimulation, cells were treated with 1% Triton X-100 and 10 mM $CaCl_2$ (final concentrations) and the light emission was recorded. To normalize, the maximal luminescence signal after histamine stimulation was divided by the total luminescence recorded per well in the entire experiment.

**Measurements of cytosolic $Ca^{2+}$.** Calibrated measurements of cytosolic $Ca^{2+}$ were made via single-cell imaging of cells loaded with 500 nM Fura-2, calibrated as previously described[25]. Population measurement of cytosolic $Ca^{2+}$ in response to 100 μM histamine was performed essentially as described in the Alliance for Cell Signalling protocols (PP00000210) except for using 4 μM Fluo-4 AM-ester and 200,000 cells per well in suspension. Minimum calibration was achieved with 5 μM ionomycin and 5 mM EGTA. Maximum calibration was achieved with 1% Triton X-100 and 20 mM $CaCl_2$ final concentrations. The $K_d$ of Fluo-4 was assumed to be 350 nM.

**Assays of $\psi_m$, oxygen consumption and NAD(P)H in intact and permeabilized cells.** For measurements of membrane potential, HeLa cells (50,000 per treatment) were seeded in eight wells of a 96-well plate and incubated with 3.5 μM JC-1 for 20 min. The uncoupler CCCP (5 μM) was added as a positive control. The wells were washed at least four times with phenol-red-free DMEM before measuring. For assaying, red fluorescence (excitation 531 nm, emission 590 nm) was measured followed by green fluorescence (excitation 488 nm, emission 535 nm) in a Perkin-Elmer Envision fluorescence plate reader. Background fluorescence values recorded from wells without cells were subtracted from all fluorescence values before calculating red/green fluorescence ratios. Intact cell respiration was measured using Seahorse XF-24 Extracellular Flux Analyzer, as previously described[26]. To investigate $Ca^{2+}$ stimulation of mitochondrial NAD(P)H in permeabilized HEK-293 cells, we adapted a previously described protocol[11].

**Measurement of mitochondrial $Ca^{2+}$ uptake in permeabilized cells.** Extramitochondrial free $Ca^{2+}$ was monitored in digitonin-permeabilized cells as previously described[5]. Calcium Green-5N fluorescence (excitation 506 nm, emission 531 nm) was monitored every 0.2 s at room temperature using a Perkin-Elmer Envision plate reader after injection of $CaCl_2$ (~50 μM final concentration). For Ru360 (Calbiochem #557440) treatments, selected concentrations of Ru360 or $H_2O$ (vehicle) were added to the digitonized cells and incubated 2 min before $CaCl_2$ injection. For calculation of relative $Ca^{2+}$ uptake, linear fits were computed for the times indicated.

**Mitochondrial isolation, mitoplast preparation, and topology analysis.** Cell lysates, crude mitochondria and mitoplasts were prepared from cultured HEK-293 cells as previously described[5]. Immunoblotting of the four fractions (5 μg) was performed with commercially available antibodies: anti-MCU (Sigma #HPA016480), anti-HSP60 (Abcam ab3080), anti-CYCS (MitoSciences #MSA06), anti-OXA1 (BD Biosciences #611980), anti-COII (MitoSciences #MS405) and anti-V5 (Invitrogen, R96025). Alkaline carbonate extraction from crude mitochondria and proteinase K digestion of mitoplasts were performed as described previously[27].

**Measurement of respiration, $\psi_m$ and $Ca^{2+}$ uptake in isolated mouse mitochondria.** Mitochondria were isolated from freshly collected mouse livers as previously described[26] and resuspended into a buffer containing 220 mM mannitol, 75 mM sucrose, 10 mM HEPES-KOH pH 7.4, 1 mM EDTA and 0.5% BSA and stored on ice until further use. Respiration and $\psi_m$ were measured optically in a well-stirred cuvette as previously described[26]. For $Ca^{2+}$ uptake measurements, 400 μg of isolated mitochondria were resuspended in 500 μl of potassium chloride (KCl) media (125 mM KCl, 2 mM $K_2HPO_4$, 1 mM $MgCl_2$, 20 mM HEPES-KOH pH 7.4) containing 1 μM Calcium Green-5N and 5 mM glutamate and 5 mM malate. Fluorescence was monitored at room temperature using a Perkin-Elmer LS-50B fluorescence spectrometer equipped with a stirring device. Respiratory control ratios (RCRs) were computed as the ratio of the ADP-stimulated respiration to the respiratory substrates-based respiration. The ADP:O ratio was computed after correcting for the oxygen leak into the cuvette. Values represent the mean ± s.d. of three, independent biological replicate mouse injections performed on three different days.

**Confocal microscopy.** HeLa cells cultured in Lab-Tek II chambered slides were co-transfected with plasmids expressing MCU–GFP and Mito-HcRed (Clontech #632434). Twenty-four hours post-transfection, cells were washed three times with PBS and imaged using a Leica TCS SP5 confocal microscope.

**Blue native studies.** Membrane protein complexes were separated by electrophoresis as previously described[28]. Briefly, mitochondria were isolated from ~20 million HeLa cells stably expressing sh-LACZ or sh-MCU according to the MitoSciences protocol (catalogue no. MS851). Total protein was quantified by bicinchoninic assay and 5 μg of mitochondria were solubilized in 1% digitonin

©2011 Macmillan Publishers Limited. All rights reserved

(w/v) on ice. Electrophoresis was performed using the NativePAGE Novex Bis-Tris Gel System from Invitrogen. Protein was transferred to a polyvinylidene fluoride (PVDF) membrane and immunoblotted with a commercially available antibody to MCU (Sigma #HPA016480). Membranes were stripped and re-probed with anti-ATP5A1 (MitoSciences MS507) as a loading control.

**Immunoprecipitation studies.** One confluent 10-cm dish of HEK-293 cells stably overexpressing either MICU1–V5 or MCU–V5 was washed twice with ice-cold PBS approximately 36 h post-transfection with either MCU–GFP or MICU1–GFP. Cells were lysed for 30 min with 1 ml of lysis buffer containing 1% *n*-dodecyl β-D-maltoside (DDM), 150 mM NaCl, and 50 mM Tris pH 7.4, and the insoluble fraction was removed by centrifugation. The soluble fraction was incubated with Dynabeads Protein G (Invitrogen 100-04D) coated with anti-V5 antibody (Invitrogen #R960-25). Immunoprecipitations were performed according to the manufacturer's guidelines. Eluents were immunoblotted with anti-V5 antibody to confirm binding and with anti-GFP antibody to probe for interactions (Novus catalogue number NB600-308). For anti-Flag immunoprecipitation studies in Supplementary Fig. 1, 2 million HEK-293 cells were plated in 10-cm plates. Cells were transfected with 2 μg of indicated plasmids using Fugene6 1 day later. After 2 days, cells were solubilized in 1 ml of lysis buffer (50 mM HEPES KOH, pH 7.4, 10 mM sodium pyrophosphate, 2 mM EDTA, 1% Triton X-100), centrifuged for 10 min at 15,000*g* and cleared lysates were incubated with anti-Flag M2 affinity gel (Sigma A2220) for 2 h at 4 °C. The immunoprecipitates were washed once with 1 ml of lysis buffer and three times with lysis buffer including 200 mM NaCl. After the last wash, the immunoprecipitates were boiled in 40 μl of SDS sample buffer. The immunoprecipitates (15 μl) were used for immunodetection using anti-Flag (Sigma F1804) and anti-MICU1 (Abcam).

©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10350

# Functional genomics reveal that the serine synthesis pathway is essential in breast cancer

Richard Possemato[1,2,3,4], Kevin M. Marks[5], Yoav D. Shaul[1,2,3,4], Michael E. Pacold[1,2,3,4,6], Dohoon Kim[1,2,3,4], Kıvanç Birsoy[1,2,3,4], Shalini Sethumadhavan[5], Hin-Koon Woo[5], Hyun G. Jang[5], Abhishek K. Jha[5], Walter W. Chen[1,2,3,4], Francesca G. Barrett[1], Nicolas Stransky[3], Zhi-Yang Tsun[1,2,3,4], Glenn S. Cowley[3], Jordi Barretina[3,7], Nada Y. Kalaany[1,2,3,4], Peggy P. Hsu[1,2,3,4], Kathleen Ottina[1,2,3,4], Albert M. Chan[1,2,3,4], Bingbing Yuan[1], Levi A. Garraway[3,7], David E. Root[3], Mari Mino-Kenudson[8], Elena F. Brachtel[8], Edward M. Driggers[5] & David M. Sabatini[1,2,3,4]

**Cancer cells adapt their metabolic processes to drive macromolecular biosynthesis for rapid cell growth and proliferation[1,2]. RNA interference (RNAi)-based loss-of-function screening has proven powerful for the identification of new and interesting cancer targets, and recent studies have used this technology *in vivo* to identify novel tumour suppressor genes[3]. Here we developed a method for identifying novel cancer targets via negative-selection RNAi screening using a human breast cancer xenograft model at an orthotopic site in the mouse. Using this method, we screened a set of metabolic genes associated with aggressive breast cancer and stemness to identify those required for *in vivo* tumorigenesis. Among the genes identified, phosphoglycerate dehydrogenase (*PHGDH*) is in a genomic region of recurrent copy number gain in breast cancer and PHGDH protein levels are elevated in 70% of oestrogen receptor (ER)-negative breast cancers. PHGDH catalyses the first step in the serine biosynthesis pathway, and breast cancer cells with high PHGDH expression have increased serine synthesis flux. Suppression of PHGDH in cell lines with elevated PHGDH expression, but not in those without, causes a strong decrease in cell proliferation and a reduction in serine synthesis. We find that PHGDH suppression does not affect intracellular serine levels, but causes a drop in the levels of α-ketoglutarate, another output of the pathway and a tricarboxylic acid (TCA) cycle intermediate. In cells with high PHGDH expression, the serine synthesis pathway contributes approximately 50% of the total anaplerotic flux of glutamine into the TCA cycle. These results reveal that certain breast cancers are dependent on increased serine pathway flux caused by PHGDH overexpression and demonstrate the utility of *in vivo* negative-selection RNAi screens for finding potential anticancer targets.**

As a starting point for identifying metabolic genes required for tumorigenesis, we cross-referenced maps of metabolic pathways with the KEGG database to compile a comprehensive list of 2,752 genes encoding all known human metabolic enzymes and transporters (Supplementary Table 1). Public oncogenomic data were analysed to score genes based on three properties: (1) higher expression in tumours versus normal tissues; (2) high expression in aggressive breast cancer; or (3) association with the stem-cell state (Fig. 1a). Genes scoring in two of these three categories as well as those at the top of each category were selected to define a high-priority set of 133 metabolic enzyme and transporter genes (Supplementary Table 2). We assembled lentiviral short hairpin RNA (shRNA) vectors targeting these genes (median 5 shRNAs per gene) and used them to generate two libraries of shRNA-expressing lentiviruses, one containing 235 distinct shRNAs (targeting transporters and control genes) and the other 516 distinct shRNAs (targeting metabolic enzymes and control genes)[4].

To identify genes that may be essential for tumorigenesis, the libraries were screened for shRNAs that become depleted during breast tumour formation in mice. Human MCF10DCIS.com cells[5] were chosen for the screens because, of several breast cancer lines examined, these were capable of forming tumours upon injection of the fewest number of cells. One and a half million MCF10DCIS.com cells were infected with each library so that each cell carried one viral integrant, and

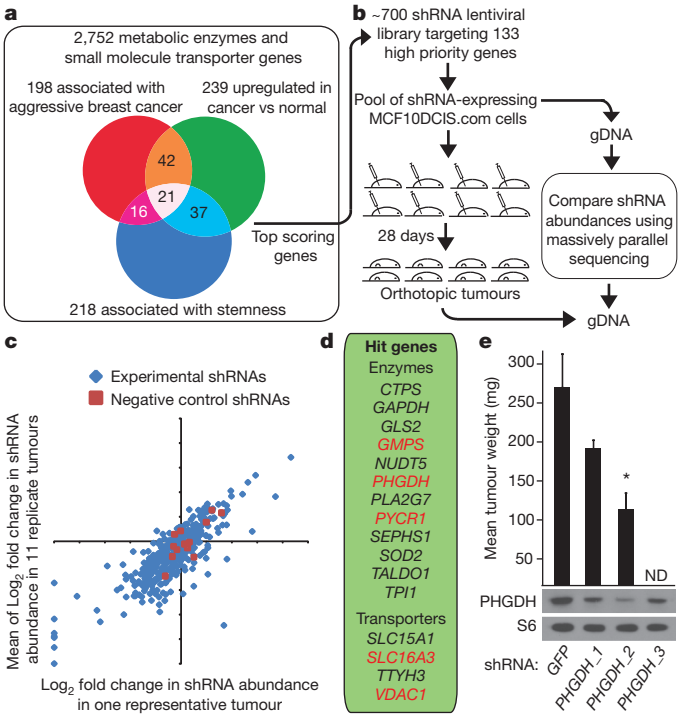



**Figure 1 | Outline of *in vivo* pooled screening strategy identifying PHGDH as essential for tumorigenesis.** **a**, Venn Diagram outlining meta-analysis. **b**, Outline of experimental design. gDNA, genomic DNA. **c**, $Log_2$ fold change in shRNAs abundance of experimental (blue) or neutral shRNAs (red) for a single tumour (*x*-axis) compared to an average of eleven tumours (*y*-axis). **d**, Genes scoring *in vivo*. **e**, Average weight of tumours from MCF10DCIS.com cells expressing shRNAs targeting PHGDH (*PHGDH_1*, *PHGDH_2* and *PHGDH_3*) or control (*GFP*) and protein expression of PHGDH or RPS6 (S6). Error bars are s.e.m. ($n = 4$). *$P$ value $< 0.05$. ND, not done.

[1]Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA. [2]Howard Hughes Medical Institute and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [3]Broad Institute of Harvard and MIT, Seven Cambridge Center, Cambridge, Massachusetts 02142, USA. [4]The David H. Koch Institute for Integrative Cancer Research at MIT, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA. [5]Agios Pharmaceuticals, 38 Sidney Street, Cambridge, Massachusetts 02139, USA. [6]Harvard Radiation Oncology Program, Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02114, USA. [7]Department of Medical Oncology and Center for Cancer Genome Discovery, Dana-Farber Cancer Institute and Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115, USA. [8]Department of Pathology, Massachusetts General Hospital and Harvard Medical School, 55 Fruit Street, Boston, Massachusetts 02114, USA.



©2011 Macmillan Publishers Limited. All rights reserved

~500–1,000 cells per shRNA (100,000–1,000,000 cells total) were injected into mouse mammary fat pads at two sites per animal (Supplementary Discussion). Twenty-eight days later orthotopic tumours were harvested and massively parallel DNA sequencing was used to determine the abundance of each shRNA in genomic DNA from tumours and initially injected cells (Fig. 1b). shRNA abundances correlated well between replicate tumours (Fig. 1c) and 5 or 12 tumours per library were analysed to identify shRNAs that became significantly depleted during tumour formation. Sixteen genes were designated hits in the screen, with at least 75% of the shRNAs targeting these genes scoring (Fig. 1d and Supplementary Table 3).

Several genes previously shown to have important roles in cancer emerged as hits, including the mitochondrial ATP transporter *VDAC1*; the lactic acid transporter *SLC16A3*; and the nucleotide synthesis genes *GMPS* and *CTPS*. The hit list also includes genes involved in the control of oxidative stress (*SOD2*, *GLS2*, *SEPHS1*), the pentose phosphate pathway (*TALDO1*), glycolysis (*GAPDH*, *TPI1*), and in the proline (*PYCR1*) and serine (*PHGDH*) biosynthetic pathways. An analogous pooled screen carried out in MCF10DCIS.com cells grown in culture rather than in tumour xenografts revealed that of 20 genes that scored in the *in vitro* screen, 10 also scored in the *in vivo* screen (Supplementary Fig. 2a, Supplementary Table 3 and Supplementary Discussion). Interestingly, *AK2*, which encodes an adenylate kinase that generates ADP from ATP and AMP, was required for *in vitro* but not *in vivo* growth (Supplementary Fig. 2b).

For five hit genes (*PHGDH*, *GMPS*, *SLC16A3*, *PYCR1* and *VDAC1*), two scoring shRNAs were tested for their effects on tumour formation. Each of these shRNAs suppressed expression of their targets in MCF10DCIS.com cells and reduced tumour-forming capacity. (Fig. 1e and Supplementary Fig. 2c). For reasons discussed later, *PHGDH* was of particular interest. The three shRNAs that scored in the *in vivo* screen also decreased PHGDH protein expression, and two shRNAs of differing knockdown efficacies inhibited tumour growth, consistent with their capacity to suppress PHGDH expression (Fig. 1e). Moreover, tumours derived from cells that in culture had confirmed reductions in PHGDH levels had, in immunohistochemical (Supplementary Fig. 3a) and immunoblotting assays (Supplementary Fig. 3b), PHGDH staining or levels similar to control tumours, suggesting that tumorigenesis selected for cells that lost shRNA-mediated PHGDH suppression.

To prioritize genes for follow-up studies we consulted a recently available analysis of copy number alterations across cancer genomes[6]. This analysis showed that *PHGDH* exists in a region of chromosome 1p commonly amplified in breast cancer and melanoma (Fig. 2a), as well as in several other cancer types (not shown). In total, 18% of patient-derived breast cancer cell lines and 6% of primary tumours have amplifications in *PHGDH*. In the data sets examined, none of the other hit genes are in genomic regions of focal and recurrent copy number gain.

Our meta-analysis for genes associated with aggressive breast cancer is corroborated by a previous study that found elevated *PHGDH* messenger RNA levels in breast cancers that are ER negative, of the basal type, and associated with poor 5-year survival[7]. We confirmed these associations in distinct gene expression data sets (Fig. 2b) and additionally found that *PHGDH* is elevated in ER-negative breast cancer relative to normal breast tissue (Fig. 2b). Of all the genes identified as hits in our screen, *PHGDH* has the most significantly elevated expression in ER-negative breast cancer (Supplementary Fig. 4). Moreover, by analysing 82 human breast tumour samples with an

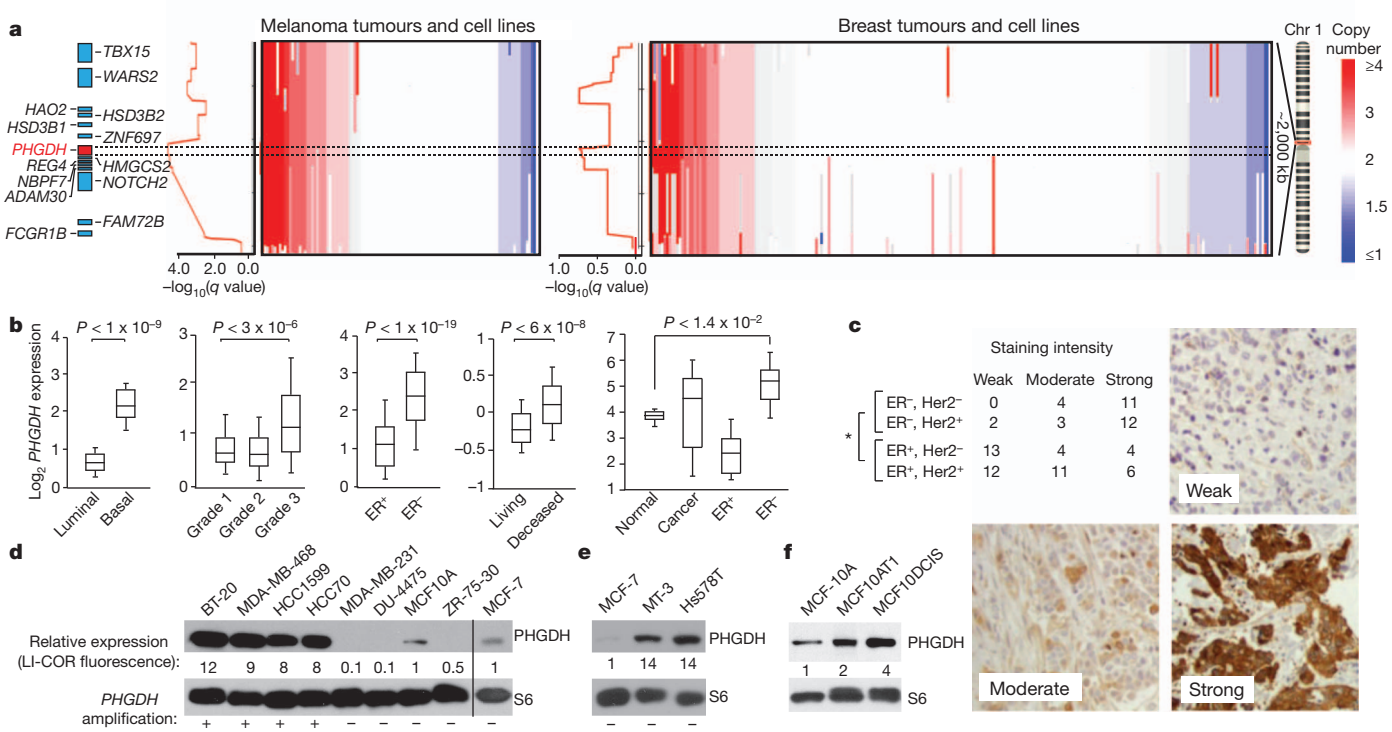



**Figure 2 | Genomic amplifications of *PHGDH* in cancer and association of PHGDH expression with aggressive breast cancer markers.** **a**, *PHGDH* vicinity copy number data for melanoma (left, $n = 111$) and breast cancer (right, $n = 243$) samples. Coloured bar indicates degree of copy number loss (blue) or gain (red). Samples sorted by copy number at *PHGDH* locus (dotted lines). Graphs at left of copy number data show amplification significance ($-\log_{10}(q$ value), ~0.60 is the significance threshold for amplification). **b**, Representative *PHGDH* gene expression data for indicated breast cancer groups. Whiskers indicate 91st and 9th percentile. **c**, Table reports numbers of human breast cancer samples with 'weak', 'moderate', or 'strong' PHGDH staining from breast cancer subgroups indicated. Representative staining intensities shown in images. Magnification, ×20. *$P < 0.0001$ comparing $ER^+$ versus $ER^-$ classes (Fisher's exact test). **d**–**f**, PHGDH protein levels are shown for *PHGDH* amplified versus non-amplified (annotated with + or −) (**d**), *PHGDH* non-amplified, over-expressing (**e**), and MCF-10A-derived cell lines (**f**). Values below PHGDH immunoblots are normalized immunoflourescent quantification (LI-COR) of PHGDH levels relative to actin control and MCF-10A and MCF-7.



©2011 Macmillan Publishers Limited. All rights reserved

immunohistochemical assay for PHGDH, we found that PHGDH protein levels correlate significantly with ER-negative status (Fig. 2c). In total, compared to ER-positive breast tumours, ~68% and ~70% of ER-negative breast tumours have elevations of PHGDH at the mRNA and protein levels, respectively (Fig. 2b, c and Supplementary Methods). ER-negative breast cancer comprises approximately 20–25% of all breast cancer cases, but as many as 50% of all breast cancer deaths within 5 years of diagnosis[8], underscoring the importance of identifying additional drug targets for this class of breast cancer.

Across a set of breast cancer lines, four lines with *PHGDH* amplifications had 8–12-fold higher PHGDH protein expression compared to non-transformed MCF-10A and ER-positive MCF-7-cell lines, which do not have *PHGDH* amplifications (Fig. 2d). Mechanisms other than gene copy number increases must also exist for boosting PHGDH expression because PHGDH protein levels were also elevated in two ER-negative cell lines (MT-3, Hs578T) lacking the *PHGDH* amplification (Fig. 2e). This is consistent with the finding that PHGDH expression is upregulated at the mRNA and protein level in a higher fraction of ER-negative breast cancers than the fraction exhibiting amplification at the DNA level. Interestingly, PHGDH is also expressed fourfold more in the MCF10DCIS.com cells used in the *in vivo* screen than in two parental lines (MCF-10A and MCF10AT) that exhibit no or lower tumorigenicity[9] (Fig. 2f).

*PHGDH* encodes 3-phosphoglycerate dehydrogenase, the first enzyme branching from glycolysis in the three-step serine biosynthetic pathway[10] (Fig. 3a). PHGDH uses NAD as a cofactor to oxidize the glycolytic intermediate 3-phosphoglycerate into phosphohydroxypyruvate[11,12], which subsequent enzymes in the pathway convert into serine via transamination (PSAT1) and phosphate ester hydrolysis (PSPH) reactions[10] (Fig. 3a). Serine is essential for synthesis of proteins and other biomolecules needed for cell proliferation, including nucleotides, phosphatidyl-serine and sphingosine (Supplementary Fig. 1). Classic studies show elevated serine biosynthetic activity, as determined by enzyme assays, in rat tumour lysates[10,13], and suggest that PSPH is the rate-limiting enzyme of this pathway in the liver[14]. Interestingly, we find that numerous genes that are expected to promote serine biosynthesis or are involved in the subsequent metabolism of serine for biosynthesis are elevated in ER-negative breast cancer (Supplementary Fig. 5), demonstrating that PHGDH elevation occurs in the context of upregulation of a broader pathway.

To understand the metabolic consequences of increased PHGDH expression we used metabolite profiling and serine synthesis pathway flux analysis to examine breast cancer cells with and without *PHGDH* amplifications. We found that cells with *PHGDH* amplifications (BT-20, MDA-MB-468 and HCC70), had increased flux through the serine synthesis pathway compared to those without *PHGDH* amplifications (MDA-MB-231, MCF7 and MCFC10A) (Fig. 3b and Supplementary Fig. 6a). Cells with elevated PHGDH and high pathway flux were capable of robust proliferation in medium lacking serine, whereas in cells with low levels of PHGDH, the deprivation of serine caused a significant blunting or even cessation of proliferation (Supplementary Fig. 6b).

PHGDH is required for the increased serine pathway flux of cells with elevated PHGDH because RNAi-mediated PHGDH suppression significantly reduced flux in MDA-MB-468 and BT-20 cells (Fig. 3c). Conversely, in MCF-10A human mammary cells engineered to overexpress PHGDH, serine pathway flux increased to levels similar to those in MDA-MB-468, BT-20 and HCC70 cells (Fig. 3d). Furthermore, MCF-10A cells overexpressing PHGDH had increased proliferation in the absence of serine, indicating that PHGDH overexpression is sufficient to drive flux through the pathway (Supplementary Fig. 6c). Interestingly, overexpression of PSPH, considered the rate-limiting serine biosynthetic enzyme in the liver, did not increase pathway flux in MCF-10A cells (Fig. 3d). The observation that PSPH is rate limiting

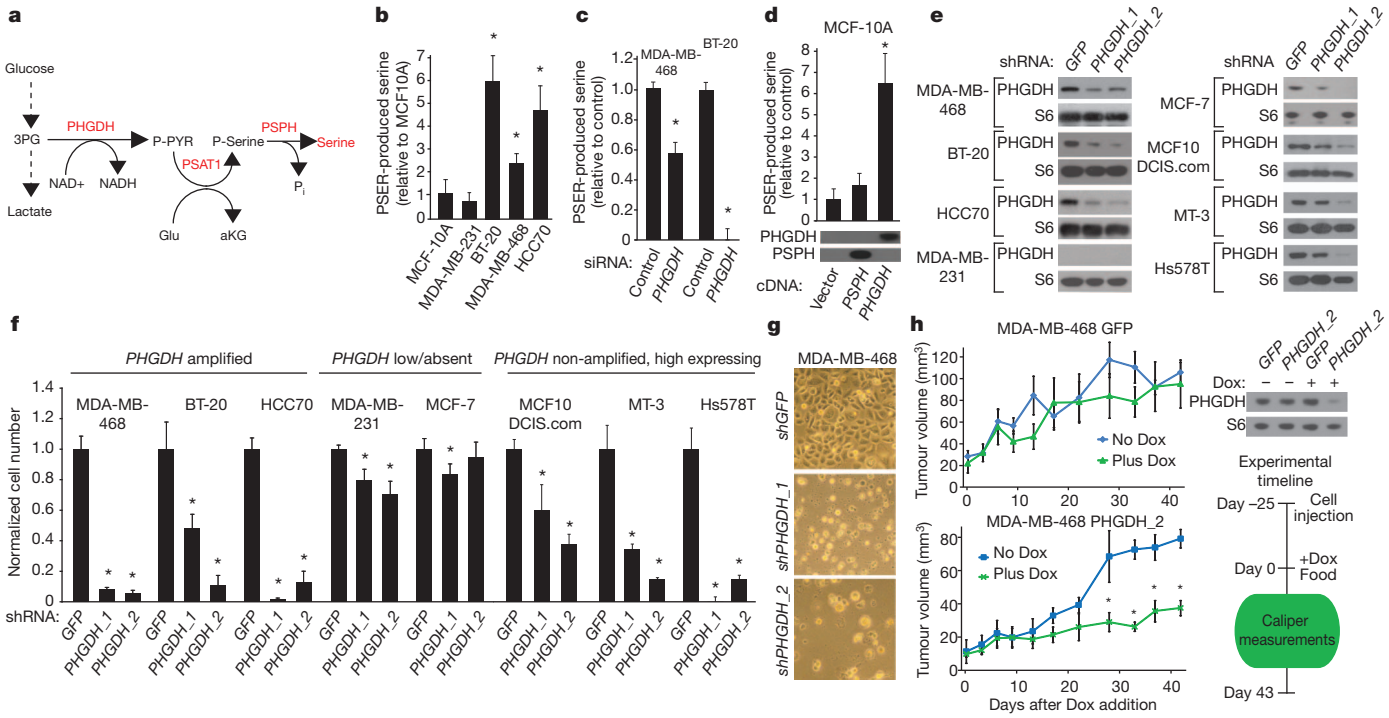



**Figure 3 | Cell lines with elevated PHGDH expression have increased serine biosynthetic pathway activity and are sensitive to PHGDH suppression.** **a**, Serine biosynthesis (PSER) pathway. **b–d**, Serine production by serine biosynthesis pathway in indicated breast cell lines (**b**), after PHGDH suppression by siRNA (**c**), and MCF-10A cells expressing *PHGDH* or *PSPH* cDNAs with associated immunoblots (**d**). **e**, Immunoblots of indicated proteins for indicated cell lines expressing control shRNA (*GFP*) or shRNAs against PHGDH (*PHGDH_1* and *PHGDH_2*). **f**, Relative proliferation of cells transduced with shRNA constructs after seven days. **g**, Images showing cellular morphology (magnification, ×20) of MDA-MB-468 at day seven of **f**. **h**, Tumour growth of MDA-MB-468 cells expressing doxycycline-inducible control shRNA (*GFP*) or shRNA against *PHGDH* (*shPHGDH_2*) in mice fed doxycycline (Dox, 2 mg kg$^{-1}$, green lines, $n = 5$) or normal (blue lines, $n = 4$) diet after initial tumour palpation (day 0). Immunoblots of PHGDH or RPS6 (S6) shown for cells *in vitro*. $^*P < 0.05$ relative to control. Error bars for metabolite measurements ($n = 4$) and tumour size indicate s.e.m., and for cell number indicate s.d. ($n = 3$).



©2011 Macmillan Publishers Limited. All rights reserved

in the liver whereas PHGDH is rate limiting in MCF-10A cells can be reconciled by the observation that serine levels in the liver (2 mM) are well above the concentration at which PSPH is feedback-inhibited by serine (500 μM), but low in cell lines in culture (~100 μM), a concentration at which PSPH should be active[14]. These data demonstrate that PHGDH is a key enzyme controlling flux through the serine biosynthetic pathway in cancer cells.

We next asked if cells with an increase in PHGDH expression require it for cell proliferation and survival. In cell lines with elevated PHGDH expression (BT-20, MDA-MB-468, HCC70, Hs578T and MT-3), but not without (MDA-MB-231 and MCF-7), RNAi-mediated suppression of PHGDH caused a marked decrease in cell number (Fig. 3e, f and Supplementary Fig. 6d) and cell death (Fig. 3g and Supplementary Fig. 6e) in the absence of apoptotic markers (Supplementary Fig. 6f). This sensitivity to PHGDH suppression was observed both in cells with *PHGDH* amplifications (BT-20, MDA-MB-468 and HCC70) and in those with high PHGDH expression but lacking *PHGDH* amplification (MT-3 and Hs578T). Consistent with flux through the serine synthesis pathway being important in cells with high PHGDH expression, suppression of the other two enzymes in the pathway (PSAT1 and PSPH) inhibited the proliferation of MDA-MB-468 and BT-20, but not MCF-7, cells (Supplementary Fig. 6g). Moreover, inhibition of PSPH inhibited tumour formation by MCF10DCIS.com cells (Supplementary Fig. 6h). Therefore, elevated PHGDH expression defines a set of breast cancer cell lines with increased serine pathway flux that are dependent upon PHGDH, PSAT1 and PSPH for proliferation. This finding suggests that many ER-negative breast cancers that express PHGDH at high levels (~70% of all ER-negative disease in our data set; Fig. 2c) may be sensitive to inhibitors of the serine synthesis pathway.

To investigate whether PHGDH suppression can affect the growth of established tumours, we generated an inducible shRNA[15] that, upon doxycycline treatment, reduced PHGDH protein levels in MDA-MB-468 cells (Fig. 3h). MDA-MB-468 cells transduced with this shRNA were allowed to form murine mammary fat pad tumours for 25 days before introduction of doxycycline in a subset of mice (Fig. 3h). Compared to control mice, those given doxycycline exhibited substantially reduced tumour growth, whereas tumours made from cells transduced with a control inducible shRNA grew equally well in the presence or absence of doxycycline (Fig. 3h). These results indicate that PHGDH suppression can adversely affect growth in existing tumours (Supplementary Discussion).

Serine is a central metabolite for biosynthetic reactions, and we find that overexpression of PHGDH contributes significantly to biosynthetic flux to serine. However, PHGDH suppression inhibited proliferation even in cells growing in media containing normal levels of extracellular serine (Fig. 3f), and supplementation with additional serine or a cell-permeable methyl-serine-ester did not blunt the effects of PHGDH suppression (Fig. 4a, b). Intracellular and extracellular serine are in equilibrium (Supplementary Fig. 7a), and import of extracellular serine was not defective in the cell lines studied (Supplementary Fig. 7b). These findings suggest that serine production may not be the only important role of PHGDH in cell lines with high PHGDH expression. We considered three hypotheses to explain our observations: (1) serine produced via the PHGDH pathway is used in a different manner than exogenous serine; (2) suppression of PHGDH adversely affects glycolysis; or (3) the PHGDH, PSAT1 and PSPH reactions produce metabolites besides serine that are also critical for cell proliferation. The first hypothesis was deemed unlikely because serine synthesized intracellularly is in equilibrium with extracellular serine (Supplementary Fig. 7a). The second hypothesis was also unlikely because PHGDH suppression did not affect glucose uptake or lactate production (Supplementary Fig. 7c).

To pursue the third hypothesis, we considered which additional metabolites the serine synthesis pathway might produce in significant levels in cells with high PHGDH expression. The serine pathway

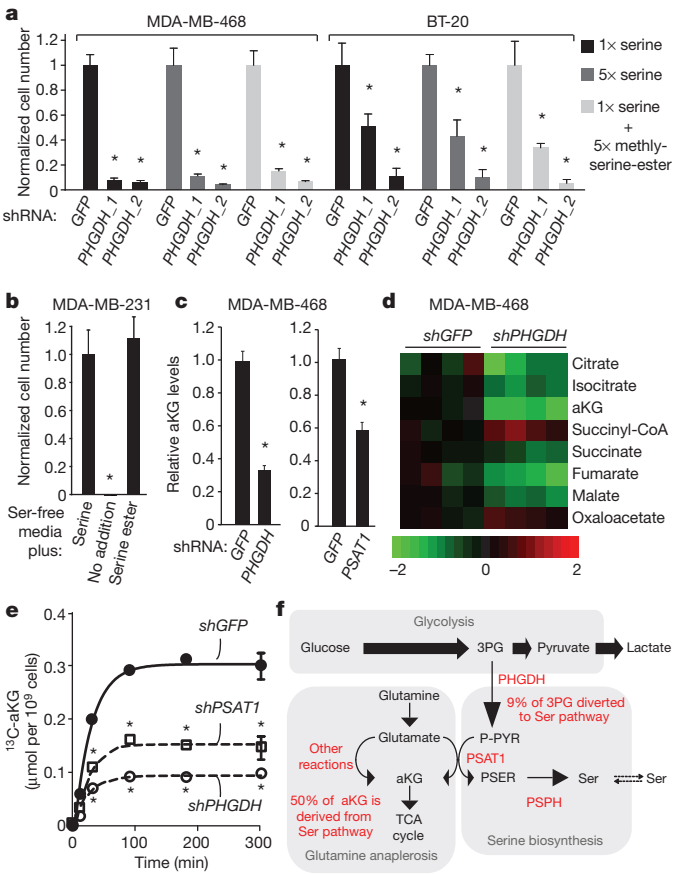



**Figure 4 | Suppression of PHGDH results in a deficiency in anaplerosis of glutamine to aKG.** **a**, Relative proliferation of cell lines indicated expressing control shRNA (*GFP*) or shRNAs against *PHGDH* (*PHGDH_1* and *PHGDH_2*) after seven days of growth under conditions indicated. **b**, Relative proliferation of MDA-MB-231 cells under conditions indicated. **c**, Intracellular aKG four days after treatment with shRNA against *PHGDH* or *PSAT1*; cell number normalized relative to control shRNA (*GFP*). **d**, TCA cycle intermediate levels four days after treatment with shRNA against *PHGDH* or *GFP* ($n = 4$). Colour bar shows $Log_2$ scale. **e**, aKG isotopic labelling at indicated time points after treatment with isotopically labelled glutamine four days after treatment with shRNA against *PHGDH*, *PSAT1* or *GFP*. **f**, Model of relative metabolite fluxes for indicated pathways. $^*P < 0.05$ relative to control. Error bars indicate s.e.m. ($n = 4$).

produces equimolar amounts of serine and α-ketoglutarate (aKG; Supplementary Fig. 1). Proliferating cells use intermediates of the TCA cycle, such as aKG, as biosynthetic precursors, and upregulate anaplerotic reactions that drive glutamine-derived carbon into the TCA cycle, counterbalancing biosynthetic efflux[16] (Supplementary Discussion). We hypothesized that in cells with high PHGDH expression, the PSAT1 reaction might contribute a significant fraction of glutamate to aKG flux. If true, the serine biosynthesis pathway would have an important role in TCA anaplerosis of glutamine-derived carbon. Consistent with this possibility, suppression of PHGDH in MDA-MB-468 cells caused a large reduction in the levels of aKG (Fig. 4c and Supplementary Fig. 7d). In fact, of the major metabolites measured, aKG was the one with the most significant and largest change upon PHGDH suppression, whereas serine levels were not significantly changed (Supplementary Fig. 8). PHGDH suppression also caused a significant reduction in other TCA components (Fig. 4d and Supplementary Fig. 8). Like suppression of PHGDH, suppression of PSAT1 also caused a significant reduction in serine pathway flux and aKG levels (Fig. 4c and Supplementary Fig. 7d, e). Furthermore, labelling studies using U-$^{13}$C-glutamine revealed that the absolute flux from glutamine to aKG and other TCA intermediates was significantly reduced in cells with RNAi-mediated suppression of PHGDH or



©2011 Macmillan Publishers Limited. All rights reserved

PSAT1 (Fig. 4e and Supplementary Fig. 9a, b). These data indicate that in cell lines with high PHGDH expression, the serine synthesis pathway is responsible for approximately 50% of the net conversion of glutamate to aKG and that suppression of PHGDH results in a significant loss of TCA intermediate flux and steady-state levels of TCA intermediates (Fig. 4f and Supplementary Fig. 9a, b). Furthermore, labelling studies using U-$^{13}$C-glucose in cell lines with *PHGDH* amplification (MDA-MB-468) and without (MDA-MB-231) revealed that in cells with high PHGDH expression, flux through the serine biosynthesis pathway shunts 8–9% of the glycolytic flux towards serine production, compared to 1–2% in the cell line with low PHGDH expression (Fig. 4f and Supplementary Fig. 9a). Therefore, increased flux through the serine biosynthesis pathway has a major impact on aKG production, but a smaller effect on glycolysis or serine availability in these cells (Supplementary Discussion). In contrast, another prominent aKG-producing transaminase, alanine aminotransferase, does not contribute significantly to aKG production in *PHGDH*-amplified cells (Supplementary Fig. 10).

We find that PHGDH expression is a critical part of a cellular program promoting serine pathway flux (Supplementary Fig. 5) and is responsible for a considerable portion of anaplerosis of glutamate into the TCA cycle as aKG (Supplementary Fig. 1). As ~70% of ER-negative breast cancers exhibit elevated PHGDH (Fig. 2c), our work suggests that targeting the serine synthesis pathway may be therapeutically valuable in breast cancers with elevated PHGDH expression or *PHGDH* amplifications (Supplementary Discussion). Lastly, we anticipate that the screening approach described here may be applicable to other cancer types or gene sets, enabling the identification of novel cancer targets directly in an *in vivo* context.

## METHODS SUMMARY

To undertake negative-selection RNAi screening in solid tumours, pools of MCF10DCIS.com cells expressing an shRNA library were injected into the fourth mammary fat pad of immunocompromised mice and allowed to form tumours. Abundances of shRNAs in the tumours were determined using massively parallel sequencing and compared to shRNA abundance in the injected cells. Genes targeted by shRNAs that were significantly depleted during tumour growth were considered hits and prioritized by analysing gene copy number data from human tumours and cancer cell lines. Lentiviral shRNAs were used to suppress PHGDH expression in breast cancer cell lines with and without *PHGDH* genomic amplification. Serine synthesis pathway activity and anaplerosis were measured via flux analyses using isotopically labelled molecules.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 19 October 2010; accepted 8 July 2011.**
**Published online 14 July 2011.**

1. Jones, R. G. & Thompson, C. B. Tumor suppressors and cell metabolism: a recipe for cancer growth. *Genes Dev.* **23,** 537–548 (2009).
2. Hsu, P. P. & Sabatini, D. M. Cancer cell metabolism: Warburg and beyond. *Cell* **134,** 703–707 (2008).
3. Bric, A. *et al.* Functional identification of tumor-suppressor genes through an *in vivo* RNA interference screen in a mouse lymphoma model. *Cancer Cell* **16,** 324–335 (2009).
4. Moffat, J. *et al.* A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124,** 1283–1298 (2006).
5. Miller, F. R., Santner, S. J., Tait, L. & Dawson, P. J. MCF10DCIS.com xenograft model of human comedo ductal carcinoma *in situ*. *J. Natl. Cancer Inst.* **92,** 1185–1186 (2000).
6. Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463,** 899–905 (2010).
7. Pollari, S. *et al.* Enhanced serine production by bone metastatic breast cancer cells stimulates osteoclastogenesis. *Breast Cancer Res. Treat.* **125,** 421–430 (2011).
8. van de Vijver, M. J. *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347,** 1999–2009 (2002).
9. Miller, F. R. *et al.* Xenograft model of progressive human proliferative breast disease. *J. Natl. Cancer Inst.* **85,** 1725–1732 (1993).
10. Snell, K. Enzymes of serine metabolism in normal, developing and neoplastic rat tissues. *Adv. Enzyme Regul.* **22,** 325–400 (1984).
11. Achouri, Y., Rider, M. H., Schaftingen, E. V. & Robbi, M. Cloning, sequencing and expression of rat liver 3-phosphoglycerate dehydrogenase. *Biochem. J.* **323,** 365–370 (1997).
12. Walsh, D. A. & Sallach, H. J. Purification and properties of chicken liver D-3-phosphoglycerate dehydrogenase. *Biochemistry* **4,** 1076–1085 (1965).
13. Knox, W. E., Herzfeld, A. & Hudson, J. Phosphoserine phosphatase distribution in normal and neoplastic rat tissues. *Arch. Biochem. Biophys.* **132,** 397–403 (1969).
14. Lund, K., Merrill, D. K. & Guynn, R. W. The reactions of the phosphorylated pathway of L-serine biosynthesis: thermodynamic relationships in rabbit liver *in vivo*. *Arch. Biochem. Biophys.* **237,** 186–196 (1985).
15. Wiederschain, D. *et al.* Single-vector inducible lentiviral RNAi system for oncology target validation. *Cell Cycle* **8,** 498–504 (2009).
16. DeBerardinis, R. J. *et al.* Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proc. Natl Acad. Sci. USA* **104,** 19345–19350 (2007).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank members of the D.M.S. laboratory, M. Kwon, B. Luo and F. Reinhardt for assistance. This research is supported by fellowships from S. G. Komen for the Cure to R.P. and the Life Science Research Foundation to Y.D.S. and grants from the Keck Foundation, the David H. Koch Institute for Integrative Cancer Research at MIT, The Alexander and Margaret Stewart Trust Fund, and NIH Grant CA103866 to D.M.S. D.M.S. is an investigator of the Howard Hughes Medical Institute.

**Author Contributions** R.P. and D.M.S. conceived the project and designed the experiments and K.M.M. and E.M.D. designed metabolic analysis and flux experiments. R.P. performed the screening, knockdown, cell proliferation and tumour formation experiments. K.B., Y.D.S., F.G.B., M.E.P., P.P.H., K.O., Z.-Y.T., N.Y.K. and W.W.C. assisted with experiments. E.F.B. collected breast cancer patient samples and scored PHGDH immunohistochemistry. D.K. performed and M.M.-K. assisted with interpretation of PHGDH immunohistochemistry. K.M.M., S.S., H.-K.W., A.K.J. and E.M.D. performed metabolite profiling, flux experiments and metabolite modelling. N.S., J.B. and L.A.G. assisted with amplification data. G.S.C. and D.E.R. assisted with screening technology. B.Y. and A.M.C. provided bioinformatic support. R.P. wrote and D.M.S. edited the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.M.S. (sabatini@wi.mit.edu).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Materials.** Materials were obtained from the following sources: antibodies to PHGDH (HPA021241) and PSPH (HPA020376) from Sigma; an antibody against PYCR1 (13108-1-AP) from Proteintech; an antibody against GMPS (A302-417A) from Bethyl Labs; an antibody against VDAC1 (ab16814) from Abcam; antibodies to RPS6 (2217), PARP (9532) and Caspase-3 (9662) from Cell Signaling Technologies; an antibody against PSAT1 (H00029968-A01) from Novus Biologicals; an antibody against SLC16A3 (AB3316P) from Millipore; and HRP-conjugated anti-mouse, anti-rabbit secondary antibodies from Santa-Cruz Biotechnology; lactate dehydrogenase from Roche (10127230001); lactic acid from Acros; RPMI-1640 media, 3-bromopyruvate and glycine/hydrazine solution (G5418) from Sigma; α-$^{15}$N-glutamine from Isotech/Sigma (486809); L-[3H(G)]-serine from Perkin Elmer; Infinity Glucose Oxidase Liquid Stable Reagent (TR15221) from Thermo Electron; U-$^{13}$C-glutamine from Isotech/Sigma (605166); MT-3 cells from DSMZ; Hs578T, MDA-MB-468, MDA-MB-231, BT-20, HCC1599, HCC70, DU4475, MCF-7 and ZR-75-30 cells from ATCC; MCF-10A, MCF-10AT1 and MCF10DCIS.com cells from the Karmanos Cancer Center, Michigan; matrigel from BD Biosciences; Phusion DNA polymerase from New England Biolabs; BCA Protein Assay from Pierce; siRNAs from Dharmacon; and amino-acid-free, glucose-free RPMI-1640 from US Biological. Lentiviral shRNAs were obtained from the The RNAi Consortium (TRC) collection of the Broad Institute[4]. The TRC numbers for the shRNAs used are: *GFP*, TRCN0000072186; *PHGDH_1*, TRCN0000221861; *PHGDH_2*, TRCN0000221865; *PSPH_1*, TRCN0000002796; *PSPH_2*, TRCN0000315168; *PSAT1_1*, TRCN0000035266; *PSAT1_2*, TRCN0000035268; *SLC16A3_1*, TRCN0000038477; *SLC16A3_2*, TRCN0000038478; *VDAC1_1*, TRCN0000029126; *VDAC1_2*, TRCN0000029127; *GMPS_1*, TRCN0000045938; *GMPS_2*, TRCN0000045941; *PYCR1_1*, TRCN0000038979; *PYCR1_2*, TRCN0000038980. The TRC website is http://www.broadinstitute.org/rnai/trc/lib. The doxycycline-inducible shRNA vector used was previously described[15].

**Cell culture.** MDA-MB-468, MDA-MB-231, BT-20, HCC1599, HCC70, DU4475, ZR-75-30, MT-3, Hs578T and MCF-7 were cultured in RPMI supplemented with 10% IFS and penicillin/streptomycin. MCF-10A and MCF10AT1 cells were cultured as described previously[17]. MCF10DCIS.com cells were cultured in 50:50 DMEM and F12 media with 5% horse serum and penicillin/streptomycin.

**Compilation of metabolic gene list.** A list of all human metabolic enzymes and small molecule transporters was generated by cross-referencing maps of metabolic pathways (Roche) with the KEGG database (http://www.genome.jp/kegg/kegg1.html). NCBI resources including Entrez Gene (http://www.ncbi.nlm.nih.gov/gene) and the available literature were used to identify known or putative gene function and to identify functional homologues. A protein was considered a metabolic enzyme if it modified a small molecule to generate another small molecule. Proteins which modified polymerized DNA or RNA or which modified proteins were excluded. In cases where an enzyme could modify both a small molecule and a macromolecule, we favoured a more liberal criterion of inclusion. A protein was considered a small molecule transporter if it formed a pore or channel through which a small molecule could traverse a lipid bilayer. Accessory or regulatory subunits of larger protein complexes were generally excluded.

**Meta-analysis of oncogenomic data.** To generate a cancer-relevant 'high priority' subset of metabolic genes (out of the 2,752 genes we classified as metabolic enzymes or small molecule transporters), we first identified those genes whose expression is significantly associated with the transformed state, advanced breast cancer, or stemness. Genes associated with the transformed state were obtained by analysing 36 gene expression studies deposited in Oncomine[18] that profiled normal human tissue and primary tumours derived from them. The gene expression profiles in each study were classified as normal or tumour and for each group the $\log_2$ median centred intensity for each gene was determined. A *P* value associated with the significance of the difference between the two groups was calculated with the Student's *t*-test. After ranking the genes based on the *P* values, the top 10% of the genes with lowest *P* values were selected from each of the 36 studies. From these genes we identified those that are in the top 10% of the most upregulated metabolic genes across the all 36 studies at a *P* value $<0.05$. Genes associated with aggressive breast cancer were obtained by analysing 15 gene expression studies from Oncomine that profiled $ER^-$ versus $ER^+$ tumours, grade 3 versus grade 1 or 2 tumours, tumours of basal versus epithelial morphology, or tumours from patients who failed to survive after 5 years of follow-up versus those who did survive at 5 years. The 15 studies were analysed as above to identify those genes that are in the top 10% of the most upregulated metabolic genes across the studies at a *P* value $<0.05$. To identify genes associated with stemness, we analysed gene expression studies comparing differentiated cells with stem cells[19], chromatin immunoprecipitation studies of stem-cell-associated transcription factors[20,21], and a previous meta-analysis of stemness-associated genes[22]. Genes were considered to be associated with stemness if their average expression was greater than fourfold upregulated in the stem versus differentiated cells profiles analysed previously[19] or if their promoters were bound by at least two stem-cell-specific transcription factors (Oct4, Nanog, Sox2, Tcf3, Dax1, Nac1 or Klf4) in both studies analysed. To generate the final high priority set of 133 genes that was screened (Supplementary Table 2), three categories of genes were selected: (1) genes scoring in all three analyses; (2) the most significantly scoring ~5% of genes in any one category; and (3) the most significantly scoring ~10% of genes in any two categories.

**Identification of cell lines for use in pooled screening.** To undertake negative-selection RNAi screening, a cell line that could form a tumour upon injection of the minimum number of cells was identified. To accomplish this, 11 breast cell lines that previously identified as capable of forming tumours were selected and 100,000 cells from each were injected into the fourth murine mammary fat pad. The cell lines tested included BT-20, BT-474, MCF10DCIS.com, HBL100, MCF7, MDA-MB-157, MDA-MB-231, MDA-MB-361, MDA-MB-453, T47D and ZR-75-1. After one month, tumours were scored by size and number scoring per site, and tumours or injection sites were analysed histologically to verify the presence of a tumour, or to identify microscopic tumours. In the timeframe of the experiment, MDA-MB-231, MDA-MB-361, MDA-MB-453, MCF7 and T47D cells formed microscopic tumours, whereas MCF10DCIS.com formed large tumours and ZR-75-1 formed small macroscopic tumours reproducibly. MCF10DCIS.com cells were then injected into murine mammary fat pads at 100,000, 10,000, 1,000 and 100 cells per site. All of these injections were capable of forming tumours, and tumour size correlated with the number of cells injected. The MCF10DCIS.com cell line was finally shown to be suitable for *in vivo* screening upon performing a screen using 180 shRNAs and demonstrating that nearly all of the shRNAs introduced initially could be recovered from the tumour and that replicate tumours exhibited significant correlation in those shRNAs over- or underrepresented compared to the injected pool. These experiments should not be construed to indicate that the excluded cell lines would not also be suitable for *in vivo* screening, as they were not tested using an shRNA pool.

**Pooled shRNA screening.** pLKO.1 lentiviral plasmids encoding shRNAs targeting the 133 transporters and metabolic enzymes listed in Supplementary Table 2 were obtained and combined to generate two plasmid pools. One contained the plasmids encoding shRNAs targeting all 47 transporters and another the plasmids encoding shRNAs targeting all 86 metabolic enzymes as well as control shRNAs designed not to target any gene. These plasmid pools were used to generate lentivirus-containing supernatants as described[23]. MCF10DCIS.com cells were infected with the pooled virus so as to ensure that each cell contained only one viral integrant. Cells were selected for 3 days with 0.5 μg ml$^{-1}$ puromycin. For the *in vivo* screen, cells were injected in 33% growth factor reduced matrigel into the fourth mammary fat pad of NOD.CB17 Scid/J mice (Jackson Labs) at 100,000 to 1,000,000 cells per injection site and tumours were harvested 4 weeks after implantation. For the *in vitro* screen, cells were plated in replicates of four at 1,000,000 per 10-cm plate and split at 1:8 once confluent (every 3–5 days) for 25–28 days. Genomic DNA was isolated from tumours or cells by digestion with proteinase K followed by isopropanol precipitation. To amplify the shRNAs encoded in the genomic DNA, PCR was performed for 33 cycles at an annealing temperature of 66 °C using 2–6 μg of genomic DNA, the primer pair indicated below, and DNA polymerase. So that PCR products obtained from many different tumours could be sequenced together, forward primers containing unique 2-nucleotide barcodes were used (see below). After purification, the PCR products from each tumour were quantified by ethidium bromide staining after gel electrophoresis, pooled at equal proportions, and analysed by high-throughput sequencing (Illumina) using the primer indicated below. shRNAs from up to 16 genomic DNA samples were sequenced together. Sequencing reads were deconvoluted using GNU Octave software by segregating the sequencing data by barcode and matching the shRNA stem sequences to those expected to be present in the shRNA pool, allowing for mismatches of up to 3 nucleotides. The $\text{Log}_2$ values reported are the average Log base 2 of the fold change in the abundance of each shRNA in the pre-injection cells compared to tumours, for $n = 5$ tumours for the transporter pool and $n = 12$ tumours for the metabolic enzyme pool, or to cells at day 25–28 for $n = 4$ *in vitro* cultures. *P* values were determined by two-sided homoscedastic unpaired *t*-test comparing each shRNA to a basket of negative-control shRNAs contained within the shRNA pools. Individual shRNAs were identified as scoring in the screens using a *P* value cutoff of 0.05 and $\text{Log}_2$ fold-change cutoff of $-1$. Genes for which >75% of the shRNAs targeting the gene scored were considered hits. Individual shRNAs were considered to be differentially required *in vitro* versus *in vivo* using a *P*-value cutoff of 0.05 by a two-sided homoscedastic unpaired *t*-test comparing the *in vitro* and *in vivo* shRNA $\text{Log}_2$ fold change scores. For the transporter pool screen, this required normalization to the median of the two distributions. shRNAs present at less than 30 reads in the pre-injection cell sample were eliminated from further analysis.

Follow-up tumour growth studies of individual genes followed a similar timeline as above, except that during PHGDH and PSPH follow-up (Fig. 1e and

©2011 Macmillan Publishers Limited. All rights reserved

Supplementary Fig. 6), 10 days elapsed between infection and injections, whereas 5 days elapsed for all other validated genes (Supplementary Fig. 2). For doxycycline-inducible constructs, MDA-MB-468 cells were infected with *GFP*- or *PHGDH*-targeting shRNAs, puromycin selected and injected into the fourth murine mammary fat pad as above. Once tumours were palpable in all animals (25 days post-injection), doxycycline chow (600 p.p.m.) was provided to a randomly assigned set of animals for the duration of the experiment. Caliper measurements were taken every 4–6 days and tumour volume was estimated by $0.5 \times W \times W \times L$, where $W$ is width and $L$ is length. All experiments involving mice were carried out with approval from the Committee for Animal Care at MIT and under supervision of the Department of Comparative Medicine at MIT.

**Primers for amplifying shRNAs encoded in genomic DNA:** Barcoded forward primer (**N** indicates location of sample-specific barcode sequence): AATGATACGGCGACCACCGAGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAA**NN**GACGAAAC. Common reverse primer: CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTTGTGGATGAATACTGCCATTTGTCTCGAGGTC. Illumina sequencing primer: AGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAA.

**Analysis of gene copy number data.** The significance of copy number alteration across multiple data sets was determined using the GISTIC algorithm with methods described in ref. 6 and using the data deposited at http://www.broadinstitute.org/tumorscape.

**Determination of proportion of tumours with *PHGDH* overexpression.** To determine the percentage of breast cancers with elevations in *PHGDH* mRNA levels, data deposited in Oncomine from ref. 8 were used. An $ER^-$ tumour was considered to have elevated *PHGDH* mRNA if the expression level was higher than 1.5 s.d. above the mean expression level in the $ER^+$ class (~91st percentile). For the percentage of breast cancer exhibiting elevated PHGDH protein, data reported in Fig. 2c were used. An $ER^-$ tumour was considered to have elevated PHGDH protein if the immunohistochemical staining signal was classified as 'high'.

**Cell proliferation assays.** For *PHGDH* or *PSPH* knockdown experiments, 10,000–20,000 MDA-MB-468, BT-20, HCC70, MCF-7, or MDA-MB-231 cells were infected with shRNA-expressing lentiviruses of known titres at a multiplicity of infection of 2.5 to 5. Cells were cultured in 12-well plates and infected via a 30-min spin at 2,250 r.p.m. in a Beckman Coulter Allegra X-12R centrifuge with an SX4750 rotor and uPlate Carrier attachment followed by an overnight incubation in media containing polybrene. Eight days after infection the number of cells was determined using a Coulter Counter (Beckman) and used to calculate relative cell proliferation. Where indicated, standard RPMI media was supplemented with serine to concentrations fivefold that of the serine already in the media. Where indicated, supplementation occurred at one and four days after lentiviral infection. For serine depletion experiments, cells were plated out as described above and the following day the standard culture medium was replaced with medium lacking serine or reconstituted with 1× serine. Dialysed serum (3 kDa MWCO) was used in serine depletion experiments except in the case of MCF-10A cells, where standard 5% serum was used.

**Immunohistochemistry and immunoblotting.** Immunoblotting was performed as described[24]. PHGDH protein levels were quantified using an Odyssey Infrared Imager (Li-Cor). For each measurement, the PHGDH signal obtained was normalized to the RPS6 signal from the same lane after accounting for background fluorescence. Immunohistochemistry was performed on formalin-fixed paraffin-embedded sections using a boiling Dako antigen retrieval method, as described[25]. A 1:250 dilution of the PHGDH antibody was used. A pathologist scored, in a blinded fashion, the intensity of the PHGDH staining in the breast tumour samples using a scale of 0–3 that represents none/weak, moderate and strong staining. Use of the tumour samples for PHGDH staining was approved by Institutional Review Boards at MIT (Protocol Number 1005003872) and Massachusetts General Hospital (Protocol Number 2010-P-001505/1).

**Glucose and lactate measurements.** Cells infected with shRNAs were plated on the day after infection at 5,000 cells per well of a 96-well plate in RPMI-10 alone or with 25 μM 3-bromopyruvate in a total of 200 μl media. On day 4 after infection, media was collected from the wells and cells were washed once with phosphate buffered saline before lysis in 50 mM NaOH. Lysate was mixed well and protein measured by BCA protein assay (Pierce). To determine the integrated protein content over the course of the assay (μg protein × days), a model was constructed with the following assumptions: control cells underwent two population doublings, cells proliferated exponentially to the final protein content, and the initial protein content for all samples was equivalent. Glucose concentration in the media was measured by glucose oxidase and peroxidase assay (Thermo Electron) and compared to control wells containing media with no cells to determine the quantity consumed. Lactate was measured by adding 5 μl of media to a solution containing 0.3 M glycine/hydrazine solution (Sigma G5418), 2.4 mM NAD+ (Fisher Scientific NC9877003), and 2 μl $ml^{-1}$ lactate dehydrogenase (5 U $\mu l^{-1}$, Roche 10127230001) in a 200 μl total volume in a 96-well microtitre plate. Plates were mixed briefly and incubated for 30 min at 37 °C before reading absorbance at 340 nm. Lactate concentration was determined by comparison to a lactic acid standard (10 mM–0 mM, Fisher Scientific AC18987-0050) and compared to control wells containing media with no cells to determine the quantity produced.

**Metabolite measurements.** For metabolite measurements, cells were cultured in cell-line-appropriate culture media (see above) in 10-cm dishes to approximately 70% confluence, typically by plating at $2 \times 10^6$ cells per dish approximately 48 h before metabolite extraction. Twenty-four hours before metabolite extraction, culture media was replenished with media containing dialysed FBS. For metabolite extraction, cells in the culture dish were rapidly washed three times with 37 °C PBS, and then metabolites were extracted by addition of 80% aqueous methanol (pre-cooled in dry-ice) followed by incubation of culture dishes on dry ice for 15 min. For quantification, a $^{13}C$-labelled internal metabolite standard for each analysed species was included in the extraction process. Cellular metabolite extracts were then collected by cell scraping and removal of the supernatant following centrifugation at 3,750 r.p.m. for 30 min (4 °C). The supernatants were then dried down using $N_2$ gas and stored dry at −80 °C before mass spectroscopy analysis. Four biological replicate samples were generated and analysed for each cell line. In addition, two parallel dishes of cells were trypsinized and counted using a Nexcelom cell counter; subsequent metabolite measurements were normalized to cell count.

All cell extracts were analysed by liquid chromatography-triple quadrapole mass spectrometry (LC-MS) using scheduled selective reaction monitoring (SRM) for each metabolite of interest, with the detector set to negative mode. Prior to injection, dried extracts were reconstituted in LCMS grade water. LC separation was achieved by the method reported[26]. Extracted metabolite concentrations were calculated from standard metabolite build-up curves using natural $^{12}C$ synthetic metabolites and normalized against cell number as well as the internal $^{13}C$-labelled metabolite standards added at the time of metabolite extraction.

**Flux analysis.** For aKG flux studies, cells were plated at 250,000 cells per well in 6-well culture dishes in typical culture media (see above). Twenty-four hours before the flux study timecourse, media was replenished with fresh RPMI media containing dialysed FBS. For the flux study timecourse, standard RPMI culture media with dialysed FBS was used and the glutamine was replaced with U-$^{13}C$ glutamine (2 mM final concentration, matching the glutamine concentration in standard RPMI culture media). At the relevant time points, metabolites were harvested as noted above.

Serine pathway flux was measured using extracellular α-$^{15}N$-glutamine, which is taken up by cells and becomes intracellular α-$^{15}N$-glutamate at a very high rate. The activity of PSAT1 (conversion of phospho-hydroxypyruvate to phosphoserine) is coupled to the transfer of the α-$^{15}N$-amino nitrogen of glutamate to phospho-hydroxypyruvate, generating aKG and α-$^{15}N$-phosphoserine. As extracellular serine is in equilibrium with the intracellular pool, the rate of accumulation of extracellular α-$^{15}N$-serine can be used to assess the activity of the serine biosynthetic pathway, and is proportional to the overall serine biosynthetic flux. For these flux studies, cells were plated at 250,000 cells per well in 6-well culture dishes in typical culture media (see above). When cells reached 60–70% confluence (typically 24–48 h post-cell plating), media was replenished with fresh media containing dialysed FBS and α-$^{15}N$-glutamine (2 mM final concentration). For the data presented in Fig. 3b, MFC10A medium was used (see above) to permit the inclusion of the MCF10A cell line, whereas for the data presented in Supplementary Fig. 6a, RPMI medium was used. Therefore, these data are not directly comparable between these two panels. Samples of media were collected from four biological replicates, at this initial time point and following 24 h of additional culture. α-$^{15}N$-serine was extracted from 300 μl of the sample media by addition of 3 volumes of acetonitrile, followed by collection of the supernatant following centrifugation for 30 min at 3,750 r.p.m. Supernatant was then dried down using $N_2$ gas and the dry samples stored at −80 °C until mass spectrometry. In parallel with the metabolite extracts, two replicate wells were trypsinized and counted at the initial time point as well as the 24 h time point. The average of these four wells was used for subsequent cell number normalization. Prior studies established the linearity of production of serine over this timecourse, and demonstrated that the intracellular and extracellular serine pools are at steady-state equilibrium, enabling measurement of a lower-bound phosphoserine pathway flux by sampling extracellular α-$^{15}N$-serine. LC-MS analysis of $^{15}N$-serine was performed using SRM in positive mode; separation was accomplished using an Atalantis HILIC Silica 5 μm (2.1 × 100 mm) column and a gradient of 10 mM of ammonium formate in Water (mobile phase A: aqueous 0.1% formic acid) and acetonitrile (mobile phase B, 0.1% formic acid) with mobile phase A linearly increasing from 5% to 60% over 4 minutes. Following a 2 min isocratic period, the system was returned to initial conditions for a total cycle time of 9 min at a flow rate of 200 μl $min^{-1}$. For flux studies, $^{13}C$-labelled internal standards were omitted in both sample extracts and standard metabolite build-up curves.

©2011 Macmillan Publishers Limited. All rights reserved

**Flux modelling.** Ordinary differential equation models were constructed for two relevant portions of central carbon metabolism, based on the schematics shown in Supplementary Fig. 9a (models (i) and (ii)). Each model consisted of 3 differential equations with the constraints of balanced flux imposed on them. These equations describe the rates of loss of unlabelled forms of metabolites after feeding of 100% U-$^{13}$C glucose or U-$^{13}$C glutamine containing media.

The fluxes were identified by minimization of an objective function to the empirical data. The choice of objective function was $\chi^2$, defined as

$$\chi^2 = \sum_{k=1}^{n} \left( \frac{(y_k - y(t_k; F))}{\sigma_k} \right)^2$$

where $y_k$ is data point $k$ with standard deviation $\sigma_k$ , and $y(t_k; F)$ is the value estimated by the model value at time point $k$ for the set of fluxes $F$. Initial fluxes before the first optimization were arbitrarily chosen as 0.1. Three independent runs of 400 fits with the trust region approach were performed, each starting from the parameter values of the currently best fit randomly disturbed by up to 4 orders of magnitude.

**Model (i).** The schematic of the upper part of glycolysis (Supplementary Fig. 9a (i)) shows that $F_2$ is the upper bound of the glycolytic flux that can be diverted to the pSer pathway. We estimated $F_2$ by fitting the model to the time course of unlabelled metabolites (3PG, PEP and lactate) obtained using LC-MS of extracts from MDA-MB-468 and MDA-MB-231 cell lines, amplified and non-amplified *PHGDH* cell lines respectively. Three independent simulations of 400 fits were run for both the cell lines. The quality of fit was characterized by $\chi^2$ values. The best 10% of the fits that also had $P$ value above a significance threshold (0.05) were chosen for the analysis. The values of the parameter had a high variability suggesting that the parameter search space resembled a shallow basin. This was confirmed by generating the $\chi^2$ landscapes for all possible pairs of parameters (data not shown). This observation suggested that additional constraints would greatly improve the predictive power of our model. Because each molecule of glucose that proceeds through glycolysis is broken into two molecules of 3PG, we imposed the requirement that $F_1$ cannot be greater than twice the measured glucose consumption rates (82 nmol per million cells per min). This additional constraint narrows down the solution of fluxes significantly, providing the results reported in the tables.

**Model (ii).** The schematic of the upstream reactions in glutaminolysis (Supplementary Fig. 9a (ii)) shows that $F_1 + F_2$ is the glutamate to aKG flux. We estimated the fluxes as described above by fitting the model to the time course of unlabelled metabolites (glutamine, glutamate and aKG) obtained using LC-MS for MDA-MB-468 cells with and without PHGDH suppression via RNAi. Identical statistical thresholds were applied as for model (i) (top 10% and $P > 0.05$) to choose solutions for the analysis. Unlike model (i), the parameters converged very well without need for further constraint, confirmed by generating the $\chi^2$ landscapes for all possible pairs of parameters (data not shown).


17. Debnath, J. & Brugge, J. S. Modelling glandular epithelial cancers in three-dimensional cultures. *Nature Rev. Cancer* **5,** 675–688 (2005).
18. Rhodes, D. R. *et al.* Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia* **9,** 166–180 (2007).
19. Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454,** 49–55 (2008).
20. Kim, J., Chu, J., Shen, X., Wang, J. & Orkin, S. H. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132,** 1049–1061 (2008).
21. Marson, A. *et al.* Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **134,** 521–533 (2008).
22. Ben-Porath, I. *et al.* An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nature Genet.* **40,** 499–507 (2008).
23. Luo, B. *et al.* Highly parallel identification of essential genes in cancer cells. *Proc. Natl Acad. Sci. USA* **105,** 20380–20385 (2008).
24. Sancak, Y. *et al.* Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell* **141,** 290–303 (2010).
25. Kalaany, N. Y. & Sabatini, D. M. Tumours with PI3K activation are resistant to dietary restriction. *Nature* **458,** 725–731 (2009).
26. Buescher, J. M., Moco, S., Sauer, U. & Zamboni, N. Ultrahigh performance liquid chromatography–tandem mass spectrometry method for fast and robust quantification of anionic and aromatic metabolites. *Anal. Chem.* **82,** 4403–4412 (2010).


©2011 **Macmillan Publishers Limited. All rights reserved**

LETTER

doi:10.1038/nature10247

# Different substrate-dependent transition states in the active site of the ribosome

Stephan Kuhlenkoetter[1], Wolfgang Wintermeyer[1] & Marina V. Rodnina[1]

**The active site of the ribosome, the peptidyl transferase centre, catalyses two reactions, namely, peptide bond formation between peptidyl-tRNA and aminoacyl-tRNA as well as the release-factor-dependent hydrolysis of peptidyl-tRNA. Unlike peptide bond formation, peptide release is strongly impaired by mutations of nucleotides within the active site, in particular by base exchanges at position A2602 (refs 1, 2). The 2′-OH group of A76 of the peptidyl-tRNA substrate seems to have a key role in peptide release[3]. According to computational analysis[4] , the 2′-OH may take part in a concerted 'proton shuttle' by which the leaving group is protonated, in analogy to similar current models of peptide bond formation[4–6]. Here we report kinetic solvent isotope effects and proton inventories (reaction rates measured in buffers with increasing content of deuterated water, $D_2O$) of the two reactions catalysed by the active site of the *Escherichia coli* ribosome. The transition state of the release factor 2 (RF2)-dependent hydrolysis reaction is characterized by the rate-limiting formation of a single strong hydrogen bond. This finding argues against a concerted proton shuttle in the transition state of the hydrolysis reaction. In comparison, the proton inventory for peptide bond formation indicates the rate-limiting formation of three hydrogen bonds with about equal contributions, consistent with a concerted eight-membered proton shuttle in the transition state[5]. Thus, the ribosome supports different rate-limiting transition states for the two reactions that take place in the peptidyl transferase centre.**

Peptide bond formation and peptide release involve the attack of different nucleophilic groups on the ester carbonyl carbon of peptidyl-transfer (t)RNA in the P site of the ribosome; these nucleophilic groups are respectively the α-amino group of A-site-bound aminoacyl-tRNA and a water molecule. Unlike peptide bond formation, which does not require auxiliary factors, peptide release is assisted by termination (release) factors; in *E. coli*, these are RF1 and RF2. On recognition of a stop codon in the decoding site, the conserved GGQ motif of RF1/2 is inserted into the peptidyl transferase centre, augmenting the active site and inducing the hydrolysis of peptidyl-tRNA (see ref. 3 for a review). Release-factor binding induces a conformational change that involves conserved 23S ribosomal RNA residues, in particular U2506 and U2585 (refs 6–10), and opens the active site for the access of water[11]. The conserved Gln residue in RF1/2 has been attributed an essential function in positioning the hydrolytic water molecule and accounts for the high specificity of the active site for water as the nucleophile, discriminating against a larger amine[12]. Computational analysis suggested that the 2′-OH of A76 of the P-site tRNA may take part in a concerted proton shuttle mechanism for the protonation of the leaving 3′-O (ref. 4), in analogy to the mechanism suggested for peptide bond formation[5,13]. A fully concerted proton shuttle would involve the formation of several (two to three) strong hydrogen bonds in the transition state. This should lead to a kinetic solvent isotope effect (KSIE) and a nonlinear proton inventory. As pertinent experimental data are lacking, we have determined the KSIE for both RF2-catalysed and uncatalysed reactions by measuring reaction rates in $H_2O$- and $D_2O$-containing buffers.

At saturation with RF2, the hydrolysis reaction in $H_2O$ proceeded with a catalytic rate constant ($k_{cat}$) of about 2 $s^{-1}$ (Fig. 1a). Reactions in $D_2O$ were found to be about four times slower than in $H_2O$ for both the RF2-catalysed reaction on the ribosome and the uncatalysed reaction (Fig. 1b). The magnitude of the KSIE indicates that it is a primary effect—that is, it arises from proton transfer(s) between groups that take part in the reaction directly[14,15]. In keeping with this contention, the same reaction rates were observed regardless of whether ribosome complexes in $D_2O$ buffer were mixed with RF2 in $H_2O$ buffer or vice versa (data not shown), establishing that the intrinsic reactivity of ribosomes and RF2 was not influenced by H–D exchange.

The proton inventory showed a linear relationship for the RF2-dependent reaction (Fig. 2a). This is consistent with a model in which the reaction rate is limited by the formation of a single strong hydrogen bond (a low-barrier hydrogen bond)[14,15]. A model involving two

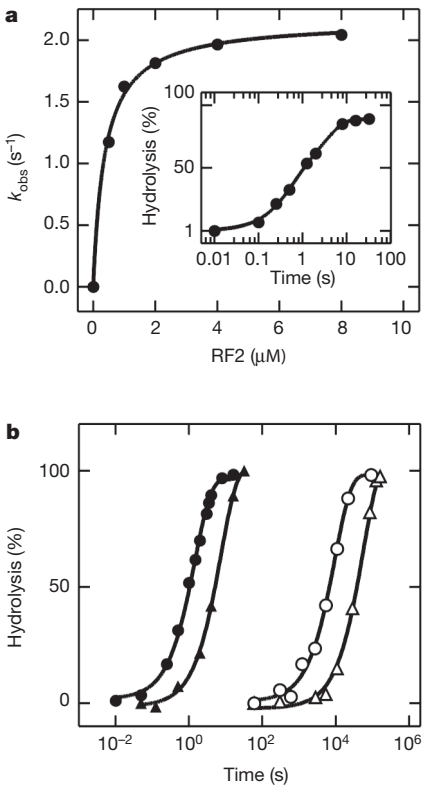



**Figure 1 | RF2-catalysed hydrolysis of fMet-tRNA$^{fMet}$ on the ribosome.** **a**, Dependence on RF2 concentration. Values of the observed rate constant ($k_{obs}$) are plotted against the concentration of RF2. The value of $k_{obs}$ measured at saturation equals the catalytic rate constant, $k_{cat}$. Inset, time course of hydrolysis (at 4 μM RF2). **b**, Time courses in $H_2O$ and $D_2O$-containing buffers. Catalysed (filled symbols) and uncatalysed (open symbols) reactions were measured in buffer with either $H_2O$ (circles) or $D_2O$ (triangles).

[1]Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany.



©2011 Macmillan Publishers Limited. All rights reserved

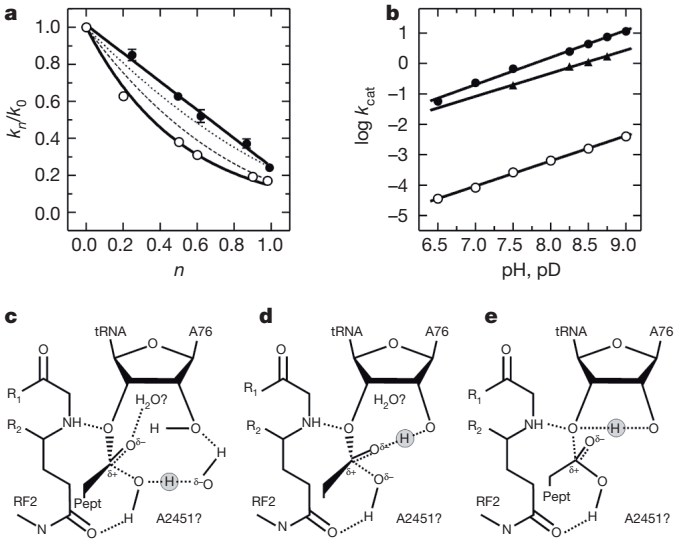



**Figure 2 | KSIE and pH dependence of the hydrolysis reaction.** **a**, Proton inventories. Time courses were measured in buffer A, pH 8.5, containing increasing mole fractions of $D_2O$, $n$, as indicated. Values of $k_{cat}$ measured in buffer containing only $H_2O$ and buffer containing $n$ mole fractions of $D_2O$ are $k_0$ and $k_n$, respectively. For the catalysed reaction (filled circles), the linear fit yields an overall KSIE of 4.1 ± 0.2; the dotted line simulates the rate-limiting simultaneous transfer of two protons in the transition-state with a fractionation factor of 0.49 each, that is, $k_n/k_0 = (1 - n + 0.49n)^2$. For the uncatalysed reaction (open circles), the curve was fitted with the function $k_n/k_0 = a^{-n}$ (ref. 14), yielding an overall KSIE of 7.0 ± 0.4; the dashed line simulates the rate-limiting transfer of three protons. Standard deviations (number of experiments ≥3) are as indicated by error bars or are smaller than the size of the symbols. **b**, pH/pD dependence. Rates of the catalysed reaction were measured at varying pH/pD in buffer B in $H_2O$ (filled circles) or $D_2O$ (filled triangles); the uncatalysed reaction was measured in buffer B in $H_2O$ (open circles). The slopes of the linear plots for the catalysed reaction were 0.89 ± 0.03 ($H_2O$) and 0.76 ± 0.03 ($D_2O$); the slope for the uncatalysed reaction was 0.83 ± 0.02 ($H_2O$). **c**–**e**, Possible transition states of the catalysed reaction (see text). The single low-barrier hydrogen bond that determines the reaction rate is encircled. The possibility of the involvement of another water molecule and/or residue A2451 of 23S rRNA is indicated.

equally contributing hydrogen bonds in the transition state was not compatible with the data. In contrast, the proton inventory of the uncatalysed hydrolysis reaction showed a pronounced downwards curvature, indicating the formation of more than one hydrogen bond in the transition state. The data for the latter reaction were fitted best with a model in which a large number of hydrogen bonds contribute about equally. This result is consistent with a model in which many water molecules form a hydrogen-bonded proton network facilitating the reaction.

We have also determined the pH/pD dependence of the RF2-dependent hydrolysis reaction (Fig. 2b). The plots of log $k_{cat}$ versus pH or pD were linear up to pH 9 with slopes of near one, suggesting that the pH/pD-dependence was determined by a single ionizing group with $pK_a > 9$ (where $K_a$ is the acid dissociation constant). The pH dependence of the uncatalysed hydrolysis exhibited about the same slope. For the latter reaction, the pH-dependence reflects the ionization of either a water molecule or the 2′-OH of the tRNA. The same groups are directly involved in the reaction on the ribosome and are, therefore, likely to contribute to the pH dependence. If the ribosome or RF2 were donating additional ionizing groups, in the simplest model the slope of the pH dependence would increase, which is not observed. Ribosomal residues in the active site (A2451, U2585 and, in particular, A2602), the replacements of which impair peptide release[1,2,16], appear to have a role in positioning the peptidyl-tRNA and/or the release factor in the active site, rather than in chemistry itself.

Substitutions of the A76 2′-OH of peptidyl-tRNA strongly inhibited peptide release by RF1, whereas the uncatalysed hydrolysis was not affected[3]. Computational analysis suggests that the 2′-OH of A76 is essential for a proton shuttle in which the 2′-OH donates a proton to the leaving group (that is, the neighbouring 3′-O of A76 of the P-site tRNA) and, at the same time, receives one from the water attacking the carbonyl carbon[4]; such a mechanism implies a six-membered transition state with two hydrogen bonds of about equal contributions and two protons synchronously changing their binding partners. The concerted proton-shuttle model is not supported by the proton inventories, which clearly show that the formation of a single strong hydrogen bond in the transition state is rate-limiting. This does not exclude the possibility that at a later, not rate-limiting, step of the reaction the leaving group is protonated via a proton shuttle.

One of several mechanisms that are consistent with our data is shown in Fig. 2c. According to this model, the attacking water molecule in the transition state is engaged in a strong hydrogen bond with a hydroxide ion that facilitates proton transfer and the nucleophilic attack. Density for a water molecule in the active site was observed in the crystal structure of a complex representing the reactant state[6]. The effect of replacing the 2′-OH group of A76 with hydrogen then may be attributed to a change of the rate-limiting step, such that proton transfer to the leaving group becomes rate-limiting. In an alternative model (Fig. 2d), the transition state may entail the attack of a hydroxide ion on the carbonyl carbon. In this case, the strong hydrogen bond could involve the 2′-OH of the P-site tRNA—which would be consistent with its important role in peptide release[3]—or a water molecule donating a proton to the carbonyl oxygen. In these models, which both imply an early transition state, the transfer of other protons, including the protonation of the leaving group, does not take place in the rate-limiting step. However, a model implying a late transition state where the protonation of the leaving group is rate-limiting would also be consistent with the data (Fig. 2e). In a late transition state, the carbonyl oxygen could be unprotonated and negatively charged (as shown in Fig. 2e) or protonated; presumably, that state would form before the rate-limiting step and therefore not contribute to the proton inventories.

Our data do not exclude additional contributions to catalysis by the Gln residue of the GGQ motif, including product and transition-state stabilization, which were suggested by crystal structures of the post-hydrolysis state[7–9]. In keeping with these suggestions, replacing the conserved Gln residue with Pro abrogated the hydrolytic activity of RF1 (ref. 9), whereas replacements with other amino acids, presenting main-chain amide groups as Gln, had only moderate effects[17].

For comparison, we have examined peptide bond formation, using the reaction with puromycin[18], a mimic of the 3′ end of aminoacyl-tRNA. The reaction exhibited a large KSIE of about eight, and the pronounced downward curvature of the proton inventory indicated the formation of more than one strong hydrogen bond in the transition state (Fig. 3a), indicating that the rate-limiting transition state of peptide bond formation differs from that of peptide release. The best fit to the data was obtained with a model in which three hydrogen bonds with about equal contributions are formed in the transition state, whereas a model assuming only two hydrogen bonds with equal contributions yielded a fit of significantly lower quality (Supplementary Fig. 2), disfavouring such a model; models assuming two or three hydrogen bonds with different contributions did not give satisfactory fits either (not shown). Thus, the proton inventory of peptide bond formation is consistent with a concerted proton shuttle in the transition state, supporting an eight-membered transition state, including the 2′-OH of the tRNA and a water molecule as proton shuttles, and three protons that move simultaneously[4,5] (Fig. 3b); the model with a six-membered transition state and only two protons 'in flight', which is not favoured by our results, is shown for comparison (Fig. 3c). Both models show early transition states, consistent with previous suggestions[19,20]. Late transition states are disfavoured by the kinetic isotope effect of the nucleophilic nitrogen and its near-zero Brønsted coefficient[19], and



©2011 Macmillan Publishers Limited. All rights reserved

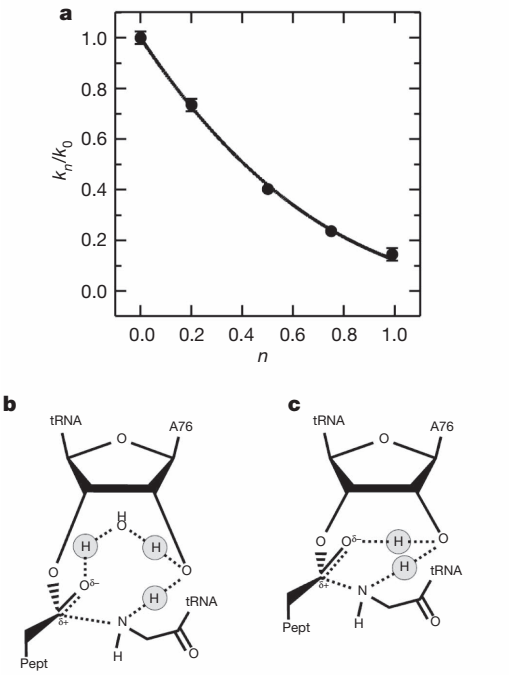



**Figure 3 | KSIE of the peptidyl transfer reaction.** **a**, Proton inventory. The puromycin (see Methods) reaction was carried out in buffer A, pH 7.5, containing increasing mole fractions of $D_2O$, $n$. The same proton inventory was obtained at pH 8.5 (not shown). The rate constant measured in $H_2O$, $k_0$, was about 0.8 $s^{-1}$, that in $D_2O$ was 0.1 $s^{-1}$. Standard deviations ($n \geq 3$) are as indicated by error bars or are smaller than the symbol size. The data were fitted best with the function $k_n/k_0 = (1 - n + 0.496n)^3$, yielding an overall KSIE of $8.2 \pm 0.2$ and the rate-limiting movement of three protons with fractionation factors of 0.496 in the transition state (Supplementary Fig. 2). **b**, **c**, Transition states of peptide bond formation with eight-membered (**b**) or six-membered (**c**) proton shuttles. Hydrogen bonds involved in the respective concerted proton shuttle[5] are encircled.

by recent work determining a full set of heavy-atom kinetic isotope effects (S. Strobel, personal communication). It is important to note that the ribosome does not provide groups that act as general acids or bases[2,21–23]. The proton shuttles as depicted result in the protonation of the carbonyl oxygen before the less favourable protonation of the leaving 3′-oxygen, following previous proposals[24,25].

The catalysis of peptide bond formation on the ribosome, compared with a spontaneous model reaction in solution, is entirely due to a change of the activation entropy, $\Delta S^{\ddagger}$ (ref. 26 and Supplementary Table 1). In contrast, the acceleration of peptide release caused by RF2 on the ribosome is due to a decrease of the activation enthalpy, $\Delta H^{\ddagger}$; the activation entropy, $\Delta S^{\ddagger}$, is not changed relative to the uncatalysed reaction (Table 1). Only in the case when the methyl group at Gln 252 in the GGQ motif of RF2 (a natural RF2 modification[27]) was lacking was there a significant entropic contribution, which may be attributed to water positioning through Gln[7].

The present results provide strong evidence indicating that the two reactions which take place in the active site of the ribosome proceed through different transition states that are distinguished by the number of protons 'in flight'. The rate-limiting transition state of the peptide release reaction is dominated by a single strong hydrogen bond. Ribosomal groups in the active site or residues of the release factor, including the conserved, functionally important GGQ motif, do not seem to take part in chemistry, but may play a role in positioning water molecules and stabilizing the transition state[4,8]. By contrast, the transition state of the peptidyl transfer reaction involves the simultaneous transfer of three protons, consistent with a concerted proton shuttle mechanism. Thus, the peptidyl transferase centre is a versatile active site, which for peptide release is augmented by residues of the release factor, providing a structured network of interactions, including ordered water molecules, that stabilizes different transition states, depending on the nature of the reactants. The catalytic mechanism employed by the ribosome illustrates how ancient RNA-world enzymes may have functioned, and how a single active site might have produced a large variety of polymers to be used as building blocks for the evolution of life.

**Table 1 | Activation parameters**

| | $E_a$ | $\Delta H^{\ddagger}$ | $\Delta G^{\ddagger}$ | $T\Delta S^{\ddagger}$ |
|---|---|---|---|---|
| Uncatalysed | 22.5 ± 0.8 | 21.9 ± 0.8 | 23.7 ± 0.6 | −1.8 ± 0.1 |
| RF2 | 16.5 ± 0.6 | 15.9 ± 0.6 | 18 ± 2 | −1.6 ± 0.2 |
| RF2* | 19.8 ± 0.7 | 19.2 ± 0.7 | 18.9 ± 0.7 | 0.3 ± 0.02 |

Activation parameters are shown for the uncatalysed and ribosome-RF2-catalysed hydrolysis reaction (kcal $mol^{-1}$). They were calculated from values of $E_a$ determined at 25 °C in buffer A, pH 7.5. RF2*, unmethylated RF2. Errors, s.d. ($n \geq 3$).

## METHODS SUMMARY

Ribosomes, initiation factors, RF2 and $tRNA^{fMet}$ were from *E. coli*. As model substrate, we have used 70S initiation complexes, that is, fMet-$tRNA^{fMet}$ bound to ribosomes programmed with an mRNA in which the AUG start codon was immediately followed by a UAA termination codon. Hydrolysis experiments were performed with RF2 that had Ala at position 246 and was methylated at N5 of Gln 252 in the GGQ motif for higher activity of the factor[27]. Peptide bond formation was studied by reacting initiation complexes with puromycin[18]. Proton inventories were measured at least three times to achieve the necessary precision[14,15].

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 27 April; accepted 3 June 2011.**
**Published online 31 July 2011.**

1. Polacek, N. *et al.* The critical role of the universally conserved A2602 of 23S ribosomal RNA in the release of the nascent peptide during translation termination. *Mol. Cell* **11,** 103–112 (2003).
2. Youngman, E. M., Brunelle, J. L., Kochaniak, A. B. & Green, R. The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release. *Cell* **117,** 589–599 (2004).
3. Brunelle, J. L., Shaw, J. J., Youngman, E. M. & Green, R. Peptide release on the ribosome depends critically on the 2' OH of the peptidyl-tRNA substrate. *RNA* **14,** 1526–1531 (2008).
4. Trobro, S. & Aqvist, J. Mechanism of the translation termination reaction on the ribosome. *Biochemistry* **48,** 11296–11303 (2009).
5. Wallin, G. & Aqvist, J. The transition state for peptide bond formation reveals the ribosome as a water trap. *Proc. Natl Acad. Sci. USA* **107,** 1888–1893 (2010).
6. Jin, H., Kelley, A. C., Loakes, D. & Ramakrishnan, V. Structure of the 70S ribosome bound to release factor 2 and a substrate analog provides insights into catalysis of peptide release. *Proc. Natl Acad. Sci. USA* **107,** 8593–8598 (2010).
7. Laurberg, M. *et al.* Structural basis for translation termination on the 70S ribosome. *Nature* **454,** 852–857 (2008).
8. Weixlbaumer, A. *et al.* Insights into translational termination from the structure of RF2 bound to the ribosome. *Science* **322,** 953–956 (2008).
9. Korostelev, A. *et al.* Crystal structure of a translation termination complex formed with release factor RF2. *Proc. Natl Acad. Sci. USA* **105,** 19684–19689 (2008).
10. Petry, S. *et al.* Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon. *Cell* **123,** 1255–1266 (2005).
11. Schmeing, T. M., Huang, K. S., Kitchen, D. E., Strobel, S. A. & Steitz, T. A. Structural insights into the roles of water and the 2' hydroxyl of the P site tRNA in the peptidyl transferase reaction. *Mol. Cell* **20,** 437–448 (2005).
12. Shaw, J. J. & Green, R. Two distinct components of release factor function uncovered by nucleophile partitioning analysis. *Mol. Cell* **28,** 458–467 (2007).
13. Trobro, S. & Aqvist, J. Analysis of predictions for the catalytic mechanism of ribosomal peptidyl transfer. *Biochemistry* **45,** 7049–7056 (2006).
14. Quinn, D. M. in *Isotope Effects in Chemistry and Biology* (eds Kohen, A. & Limbach, H. H.) 995–1018 (CRC Taylor and Francis, 2006).
15. Schowen, K. B. J. in *Transition States of Biochemical Processes* (eds Gandour, R. D. & Schowen, R. L.) 225–283 (Plenum, 1972).
16. Amort, M. *et al.* An intact ribose moiety at A2602 of 23S rRNA is key to trigger peptidyl-tRNA hydrolysis during translation termination. *Nucleic Acids Res.* **35,** 5130–5140 (2007).
17. Mora, L. *et al.* The essential role of the invariant GGQ motif in the function and stability *in vivo* of bacterial release factors RF1 and RF2. *Mol. Microbiol.* **47,** 267–275 (2003).
18. Katunin, V. I., Muth, G. W., Strobel, S. A., Wintermeyer, W. & Rodnina, M. V. Important contribution to catalysis of peptide bond formation by a single ionizing group within the ribosome. *Mol. Cell* **10,** 339–346 (2002).



©2011 Macmillan Publishers Limited. All rights reserved

19. Kingery, D. A. *et al.* An uncharged amine in the transition state of the ribosomal peptidyl transfer reaction. *Chem. Biol.* **15,** 493–500 (2008).
20. Seila, A. C., Okuda, K., Nunez, S., Seila, A. F. & Strobel, S. A. Kinetic isotope effect analysis of the ribosomal peptidyl transferase reaction. *Biochemistry* **44,** 4018–4027 (2005).
21. Beringer, M., Adio, S., Wintermeyer, W. & Rodnina, M. The G2447A mutation does not affect ionization of a ribosomal group taking part in peptide bond formation. *RNA* **9,** 919–922 (2003).
22. Beringer, M. *et al.* Essential mechanisms in the catalysis of peptide bond formation on the ribosome. *J. Biol. Chem.* **280,** 36065–36072 (2005).
23. Bieling, P., Beringer, M., Adio, S. & Rodnina, M. V. Peptide bond formation does not involve acid-base catalysis by ribosomal residues. *Nature Struct. Mol. Biol.* **13,** 423–428 (2006).
24. Huang, K. S., Carrasco, N., Pfund, E. & Strobel, S. A. Transition state chirality and role of the vicinal hydroxyl in the ribosomal peptidyl transferase reaction. *Biochemistry* **47,** 8822–8827 (2008).
25. Rangelov, M. A., Vayssilov, G. N., Yomtova, V. M. & Petkov, D. D. The syn-oriented 2-OH provides a favorable proton transfer geometry in 1,2-diol monoester aminolysis: implications for the ribosome mechanism. *J. Am. Chem. Soc.* **128,** 4964–4965 (2006).
26. Sievers, A., Beringer, M., Rodnina, M. V. & Wolfenden, R. The ribosome as an entropy trap. *Proc. Natl Acad. Sci. USA* **101,** 7897–7901 (2004).
27. Dincbas-Renqvist, V. *et al.* A post-translational modification in the GGQ motif of RF2 from *Escherichia coli* stimulates termination of translation. *EMBO J.* **19,** 6900–6907 (2000).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank H.-H. Limbach for discussions and advice. This work was supported by the Deutsche Forschungsgemeinschaft (M.V.R. and W.W.).

**Author Contributions** S.K., W.W. and M.V.R. conceived the study and designed experiments. S.K. performed experiments. All three authors discussed results and wrote the paper.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.V.R. (rodnina@mpibpc.mpg.de).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Experimental system.** As model substrate for peptide release we have used fMet-tRNA$^{fMet}$ bound to ribosomes programmed with an mRNA in which the AUG start codon was immediately followed by a UAA termination codon[12]. Hydrolysis experiments were performed with RF2 that had Ala at position 246 and that was fully methylated at N5 of Gln252 *in vitro* for higher activity of the factor[27]. Measurements were performed at saturating concentrations of RF2 (4 μM), where the binding of the factor does not limit the reaction rate. The $k_{cat}$ value for the RF2-dependent hydrolysis of fMet-tRNA$^{fMet}$ that we obtained (2.2 $s^{-1}$; Fig. 1a) is comparable with published $k_{cat}$ values for the hydrolysis of peptidyl-tRNA carrying peptide chains of four amino acids (1.5 $s^{-1}$, ref. 28; ~6 $s^{-1}$, ref. 27). As model substrate for peptide bond formation, 70S initiation complexes were used that were programmed with a similar mRNA, except that a UUC codon followed the AUG codon. Time courses were measured with saturating concentration of puromycin (10 mM), that is, at conditions where the chemistry step is rate-limiting[18,26].

**Materials.** Ribosomes from *E. coli* MRE600, initiation factors, and f[$^3$H]Met-tRNA$^{fMet}$ were prepared as described[18,21–23]. The mRNAs (5′-GGCAAGGAGGUAAAUAAUGUAAACGAUU-3′, initiation and stop codons underlined, or a similar mRNA with a UUC codon following the initiation codon) were purchased from CureVac. Deuterium oxide (99.95%) was obtained from Deutero GmbH. $^3$H-labelled L-methionine and [methyl-$^3$H]S-adenosyl-L-methionine were from Hartmann Analytics.

**Initiation complexes.** Ribosomes (1.5 μM), mRNA (3 μM), f[$^3$H]Met-tRNA$^{fMet}$ (2.7 μM), initiation factors 1, 2 and 3 (1.8 μM each) and GTP (1 mM) in a total volume of 6.4 ml were incubated for 30 min at 37 °C. Complexes were subsequently purified by ultracentrifugation, where 1.6 ml of the reaction mixture were centrifuged through a 400 μl sucrose cushion (20 mM HEPES-HCl, pH 7.5, 100 mM KCl, 20 mM $MgCl_2$, 40% sucrose). Pellets were resuspended in the appropriate buffer (see below). The initiation efficiency was controlled by nitrocellulose filtration and was >95% for the complexes used for the experiments. Filters were dissolved in 10 ml of scintillation liquid (Quickszint 361; Zinsser Analytic), and the radioactivity was determined by scintillation counting (TriCarb, Perkin Elmer).

**Preparation of RF2.** The RF2 gene was amplified by PCR from *E. coli* JM109 genomic DNA (Ala at position 246) and cloned into a pET30a vector. The T of the inner in-frame TAG stop codon of the RF2 reading frame was deleted by the quick-change method, eliminating the necessity of the +1 frameshift for RF2 expression. *E. coli* Bl21(DE3) cells were transformed with a pET30a/RF2 construct coding for RF2 extended by a C-terminal six-histidine tag. The protein was purified by $Ni^{2+}$-NTA affinity chromatography and on a HiTrap Q FF (GE Healthcare) column. The His-tag did not affect the functional activity of RF2, as verified using an untagged protein. If not stated otherwise, RF2 was $N^5$-methylated at Gln252. For methylation *in vitro*, RF2 (10 μM), S-adenosyl-L-methionine (50 μM) and the methyl transferase HemK (0.5 μM; also designated PrmC) were incubated overnight at room temperature in a total volume of 18 ml in buffer A. Methylated RF2 was purified by gel filtration on a HiLoad 26/60 Superdex 75 column (GE Healthcare). The extent of methylation was close to 100%, as analysed by filter binding of RF2 that was radioactively labelled with [methyl-$^3$H]S-adenosyl-L-methionine or by HPLC ESI-MS.

**Kinetic measurements.** Experiments were carried out at 37 °C in buffer A (20 mM HEPES-HCl, 100 mM KCl, 7 mM $MgCl_2$) adjusted to pH 7.5 or 8.5, as indicated. For measuring pH/pD dependencies, buffer B (50 mM Tris-HCl, 20 mM Bis-Tris-HCl, 70 mM $NH_4Cl$, 30 mM KCl, 7 mM $MgCl_2$) was used. Time courses of RF2-catalysed f[$^3$H]Met-tRNA$^{fMet}$ hydrolysis on the ribosome were started by rapidly mixing equal volumes (14 μl) of initiation complex (0.5 μM) and RF2 (8 μM) in a quench-flow apparatus (KinTek Corp). Reactions were terminated by TCA precipitation, and remaining intact f[$^3$H]Met-tRNA$^{fMet}$ was collected on nitrocellulose filters and quantified by scintillation counting. Time courses of the uncatalysed reaction, measured with f[$^3$H]Met-tRNA$^{fMet}$ free in solution, were started by manual mixing. For determining the pH dependence, it was verified that the concentration of RF2 used was saturating over the whole pH range (data not shown). Time courses of peptide bond formation were measured in buffer A (pH 7.5 and 8.5) by rapidly mixing equal volumes of the respective initiation complexes (0.5 μM) and puromycin (20 mM). The reaction was stopped and the remaining intact f[$^3$H]Met-tRNA$^{fMet}$ quantified as described above.

**Determination of activation parameters.** Activation energies ($E_a$) were determined from the slopes of the Arrhenius plots (Supplementary Fig. 1), which yielded $-E_a/R$. The enthalpy of activation was calculated as $\Delta H^{\ddagger} = E_a - RT$, where $R$ is the gas constant (1.9872 cal $K^{-1}$ $mol^{-1}$) and $T$ the temperature (298.15°K). Free energies of activation were calculated from $\Delta G^{\ddagger} = -RT \ln[(k_{cat}h)/(k_BT)]$, where $h$ is Planck's constant and $k_B$ is Boltzmann's constant. $T\Delta S^{\ddagger}$ was calculated as the difference between $\Delta G^{\ddagger}$ and $\Delta H^{\ddagger}$.

**Determination of kinetic solvent isotope effects.** The KSIE was determined by comparing reaction rates in $H_2O$- and $D_2O$-containing buffer. Buffers were prepared with $D_2O$, and the pD (negative logarithm of [$D^+$]) was adjusted taking into account the 0.4 unit shift for glass electrode measurements in $D_2O$ (ref. 29), which ensures equal concentrations of $OD^-$ and $OH^-$ in $D_2O$ and $H_2O$, respectively. Initiation complexes were prepared and purified in buffers with the desired $H_2O$/$D_2O$ ratio. RF2 was dialysed overnight at 4 °C into the $D_2O$-containing buffers using dialysis cassettes (10 kDa molecular weight cut-off; Pierce). Proton inventories were measured at least three times to achieve the necessary precision[14,15].

28. Freistroffer, D. V., Kwiatkowski, M., Buckingham, R. H. & Ehrenberg, M. The accuracy of codon recognition by polypeptide release factors. *Proc. Natl Acad. Sci. USA* **97,** 2046–2051 (2000).
29. Glasoe, P. K. & Long, F. A. Use of glass electrodes to measure acidities in deuterium oxide. *J. Phys. Chem.* **64,** 188–190 (1960).

©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10333

# Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals

Clementina Dellomonaco[1], James M. Clomburg[1], Elliot N. Miller[1] & Ramon Gonzalez[1,2]

**Advanced (long-chain) fuels and chemicals are generated from short-chain metabolic intermediates through pathways that require carbon-chain elongation. The condensation reactions mediating this carbon–carbon bond formation can be catalysed by enzymes from the thiolase superfamily, including β-ketoacyl-acyl-carrier protein (ACP) synthases, polyketide synthases, 3-hydroxy-3-methylglutaryl-CoA synthases, and biosynthetic thiolases[1]. Pathways involving these enzymes have been exploited for fuel and chemical production, with fatty-acid biosynthesis (β-ketoacyl-ACP synthases) attracting the most attention in recent years[2–4]. Degradative thiolases, which are part of the thiolase superfamily and naturally function in the β-oxidation of fatty acids[5,6], can also operate in the synthetic direction and thus enable carbon-chain elongation. Here we demonstrate that a functional reversal of the β-oxidation cycle can be used as a metabolic platform for the synthesis of alcohols and carboxylic acids with various chain lengths and functionalities. This pathway operates with coenzyme A (CoA) thioester intermediates and directly uses acetyl-CoA for acyl-chain elongation (rather than first requiring ATP-dependent activation to malonyl-CoA), characteristics that enable product synthesis at maximum carbon and energy efficiency. The reversal of the β-oxidation cycle was engineered in *Escherichia coli* and used in combination with endogenous dehydrogenases and thioesterases to synthesize *n*-alcohols, fatty acids and 3-hydroxy-, 3-keto- and trans-$\Delta^2$-carboxylic acids. The superior nature of the engineered pathway was demonstrated by producing higher-chain linear *n*-alcohols ($C \geq 4$) and extracellular long-chain fatty acids ($C > 10$) at higher efficiency than previously reported[2,4,7–9]. The ubiquitous nature of β-oxidation, aldehyde/alcohol dehydrogenase and thioesterase enzymes has the potential to enable the efficient synthesis of these products in other industrial organisms.**

Engineering the β-oxidation cycle as a metabolic platform for the synthesis of advanced fuels and chemicals (Fig. 1) requires the reverse

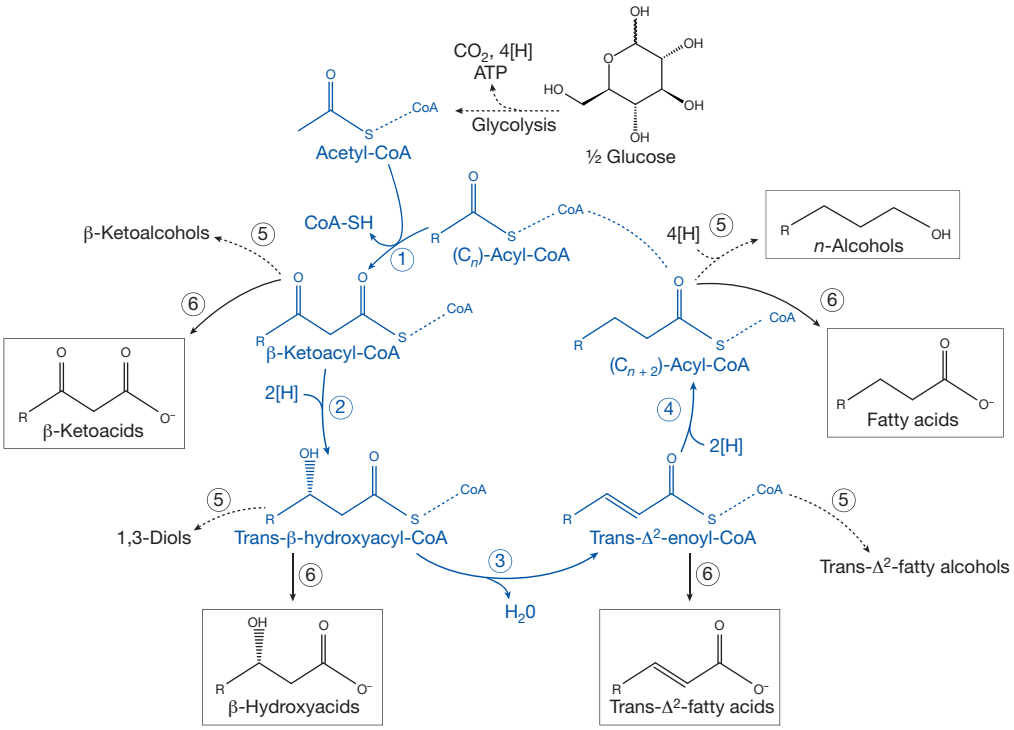



**Figure 1 | Functional reversal of the β-oxidation cycle.** The following enzymes (gene names in parentheses) are involved in the functional reversal of the β-oxidation cycle engineered in this work (shown in blue): (1) thiolase (*yqeF*, *fadA*); (2) hydroxyacyl-CoA dehydrogenase (*fadB*); (3) enoyl-CoA hydratase (*fadB*); (4) enoyl-CoA reductase (*ydiO*). Each turn of the cycle generates an acyl-CoA that is two carbons longer than the initial acyl-CoA thioester (indicated as $C_{n+2}$). Intermediates in the engineered pathway can be converted to a functionally diverse set of molecules using aldehyde-forming acyl-CoA reductases and alcohol dehydrogenases (5) and acyl-CoA thioesterases (6). Products for which synthesis was demonstrated in this study are boxed. R, side chain attached to the acyl-CoA group of the starter molecule (for example, R = H for acetyl-CoA and R = $CH_3$ for propionyl-CoA); 2[H], NADH/NADPH/ferredoxin. Dotted lines indicate multiple steps. Dashed lines without arrowheads connect identical metabolites of different chain length. A comparison of *n*-alcohol production through the engineered β-oxidation cycle and the recently proposed fatty-acid biosynthesis pathway[2] is shown in Supplementary Fig. 1.

[1]Department of Chemical and Biomolecular Engineering, Rice University, Houston, Texas 77005, USA. [2]Department of Bioengineering, Rice University, Houston, Texas 77005, USA.



©2011 Macmillan Publishers Limited. All rights reserved

operation of this pathway in the absence of its natural substrate (that is, fatty acids) and presence of a non-fatty-acid carbon source (for example, glucose). Constitutive expression of the *fad* and *ato* regulons, which encode the β-oxidation system in *Escherichia coli*[5], in the absence of fatty acids was achieved by introducing previously reported *fadR* and *atoC(c)* mutations[10] (Supplementary Table 1). Several operons encoding β-oxidation cycle enzymes are also activated by the cyclic AMP (cAMP) receptor protein (CRP)–cAMP complex and therefore subjected to carbon catabolite repression in the presence of glucose[11]. This regulatory mechanism was circumvented by replacing the native *crp* gene with a cAMP-independent mutant (*crp**) previously reported to confer a catabolite-derepressed phenotype[12] (Supplementary Table 1). Finally, the *arcA* gene was deleted to prevent ArcA-mediated repression of most operons encoding the β-oxidation cycle[13].

The genetic modifications outlined above were introduced into wild-type *E. coli* MG1655 (resulting strain named RB01: MG1655 *fadR atoC(c) crp** Δ*arc*A) and in a strain in which other fermentative pathways were blocked (resulting strain named RB02: MG1655 *fadR atoC(c) crp** Δ*arc*A Δ*adhE* Δ*pta* Δ*frd*A). However, none of the products that can be derived from the β-oxidation cycle intermediates (Fig. 1) was synthesized by these strains (Supplementary Table 2). A limitation on the expression of β-oxidation enzymes was ruled out after confirming the presence of these activities in strain RB02 (Table 1a).

The expression of appropriate termination enzymes is vital to achieve the conversion of the CoA-thioester intermediates in the reversal of the β-oxidation cycle to the desired product(s) (Fig. 1). We first targeted the synthesis of *n*-butanol, as this product can be generated in a one-turn reversal of the β-oxidation cycle. Unlike the significant activity observed for key β-oxidation enzymes, the level of *n*-butanol dehydrogenase in strain RB02 was very low (Table 1a). To address this issue, two endogenous aldehyde/alcohol dehydrogenases with high sequence and structural similarity to clostridial butyraldehyde/butanol dehydrogenase were expressed in strain RB02: that is, L-1,2-propanediol oxidoreductase (FucO) and an aldehyde/alcohol dehydrogenase (YqhD) (Supplementary Table 3). These modifications enabled a functional one-turn reversal of the β-oxidation cycle, as evidenced by the synthesis of *n*-butanol in strains RB02 [*yqhD*+] and RB02 [*fucO*+] (Fig. 2a). Despite the potential for YqhD to catalyse the conversion of butyraldehyde to *n*-butanol[14], overexpression of FucO led to higher concentration of *n*-butanol and lower concentration of the major fermentation by-product, ethanol (Fig. 2a).

The initiation of the reversal of the β-oxidation cycle (that is, conversion of acetyl-CoA to acetoacetyl-CoA) is another key step in the engineered pathway. Therefore, we evaluated the effect of overexpressing *atoB*-encoded acetyl-CoA acetyltransferase, which has high specificity for short-chain acyl-CoA molecules[15,16] and a predicted acyltransferase (YqeF) with high sequence similarity to AtoB. These modifications also supported a functional reversal of the β-oxidation cycle, as evidenced by the synthesis of *n*-butanol in strains RB02 [*atoB*+] and RB02 [*yqeF*+] (Fig. 2a). The simultaneous overexpression of YqeF and FucO in strain RB02 yielded a high *n*-butanol titre (1.9 g $l^{-1}$), *n*-butanol yield (0.19 g *n*-butanol per g total glucose consumed) and *n*-butanol-to-ethanol ratio (>5:1) (Fig. 2a). No *n*-butanol was synthesized upon overexpression of FucO and YqeF in wild-type MG1655 (Supplementary Table 2), demonstrating the requirement of an active reversal of the β-oxidation cycle.

Further reduction in the synthesis of by-product ethanol was realized through deletion of *eutE* and *yqhD*: EutE is an aldehyde dehydrogenase with high sequence similarity to AdhE and YqhD overexpression favoured the synthesis of ethanol over *n*-butanol (Fig. 2a). The resulting strain (RB02 Δ*yqhD* Δ*eutE* [*yqeF*+*fucO*+]) synthesized 2.2 g $l^{-1}$ of *n*-butanol in 24 h at a yield of 0.28 g *n*-butanol per g total glucose consumed (Fig. 2a). When grown in a bioreactor using a higher concentration of glucose, this strain produced *n*-butanol at high titre

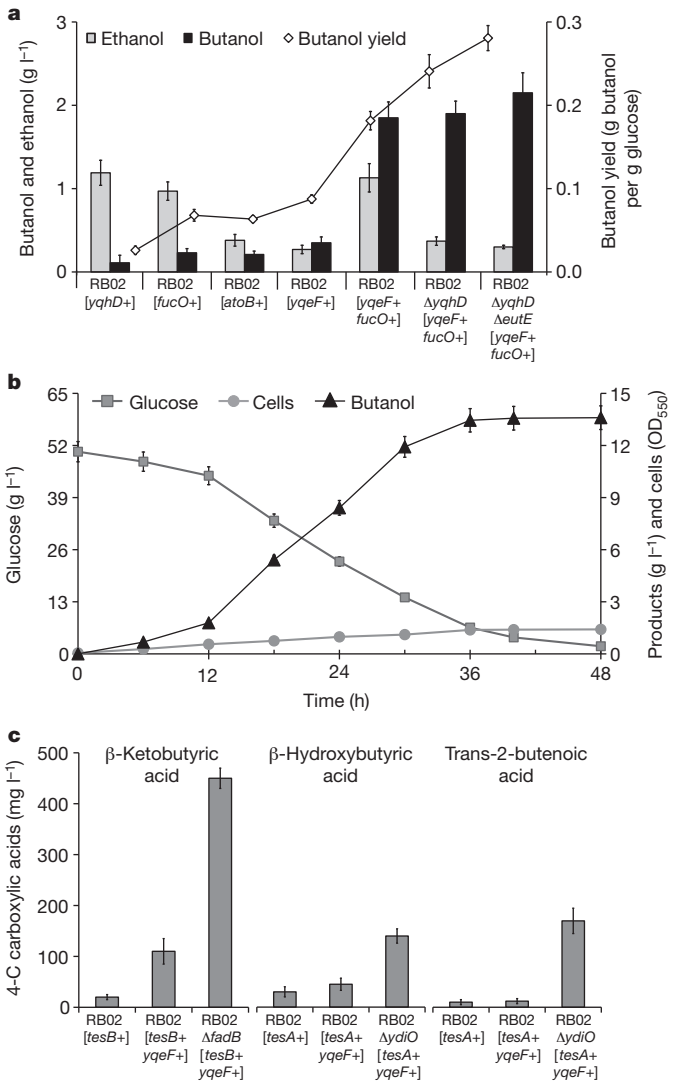



**Figure 2 | Engineered one-turn reversal of the β-oxidation cycle for the synthesis of *n*-butanol and short-chain carboxylic acids.** **a**, Effect of gene overexpression and knockout (indicated below *x* axis) on the synthesis of *n*-butanol and ethanol in strain RB02 (*fadR atoC(c) crp** Δ*arcA* Δ*adhE* Δ*pta* Δ*frdA*). Experiments were performed at 30 °C for 24 h in shake flasks using glucose (2% w/v) minimal medium. The *n*-butanol yield was calculated as g *n*-butanol per g total glucose consumed. **b**, Kinetics of *n*-butanol production by strain RB02 Δ*yqhD* Δ*eutE* [*yqeF*+*fucO*+]. Cells were cultivated in fermentors containing minimal medium supplemented with 5% (w/v) glucose. The dissolved oxygen was controlled at 5% of saturation, temperature at 30 °C, and pH at 7. **c**, Synthesis of β-ketobutyric acid (left), β-hydroxybutyric acid (centre) and trans-2-butenoic acid (right) upon overexpression of thioesterases I (TesA) and II (TesB) in strains RB02, RB02 Δ*fadB* and RB02 Δ*ydiO*. Experiments were run at 37 °C for 48 h in shake flasks using glucose (1% w/v) minimal medium. Error bars represent the standard deviation of three replicates.

(~14 g $l^{-1}$), yield (0.33 g *n*-butanol per g total glucose consumed) and rate (~2 g *n*-butanol per g cell dry weight per h) (Fig. 2b). This performance, which was achieved in the absence of rich nutrients and without importing foreign genes, is better than reported for other organisms engineered for *n*-butanol production (Supplementary Table 4) and also surpasses the *n*-butanol yield and specific productivity reported for native producers[17]. The engineered reversal of the β-oxidation cycle operated at a maximum flux of 73.4 mmol acetyl-CoA per g cell dry weight per h (12–18 h in Fig. 2b), which exceeds the flux reported in the literature for native or engineered fermentative pathways[18,19].

A characterization of the engineered β-oxidation pathway was conducted to confirm the identity of the enzymes catalysing each



©2011 Macmillan Publishers Limited. All rights reserved

individual step and their role in *n*-butanol synthesis (Table 1). Enzyme activity measurements showed a high level of expression of key pathway enzymes in strain RB02 [*yqeF*+*fucO*+] but negligible activity in wild-type MG1655 (Table 1a). Gene knockout and gene complementation experiments along with quantification of fermentation products (Table 1b) demonstrated that the primary enzymes involved in the synthesis of *n*-butanol through the engineered reversal of the β-oxidation pathway are (activities shown in parentheses along with reaction numbers, as per Fig. 1): YqeF (predicted acyltransferase: reaction (1)); FadB (3-hydroxyacyl-CoA dehydrogenase and enoyl-CoA hydratase: reactions (2) and (3), respectively); YdiO (predicted acyl-CoA dehydrogenase/enoyl-CoA reductase: reaction (4)), and MhpF and FucO (aldehyde-forming acyl-CoA reductase and *n*-butanol dehydrogenase, respectively: reaction (5)).

YdiO is proposed to catalyse the reduction of enoyl-CoA to acyl-CoA (reaction (4)). The reverse of this reaction is catalysed by FadE and is the only irreversible step in the catabolic operation of the β-oxidation cycle[5]. In agreement with our proposal, deletion of *ydiO* in strain RB02 [*yqeF*+*fucO*+] completely abolished *n*-butanol synthesis (Table 1b). Although *ydiO* was previously proposed to encode an acyl-CoA dehydrogenase that would replace FadE during the anaerobic catabolism of fatty acids[20], a sequence comparison between YdiO and *E. coli* proteins does not reveal a significant similarity to FadE (Supplementary Table 5). In contrast, YdiO shares high homology with crotonobetainyl-CoA reductase (CaiA; Supplementary Table 5). CaiA catalyses the reduction of crotonobetainyl-CoA to γ-butyrobetainyl-CoA[21], a reaction similar to that catalysed by YdiO in the reversal of the β-oxidation cycle. Moreover, the operon *fixABCX* is required for the transfer of electrons to CaiA[22,23] and encodes flavoproteins and ferredoxin with high sequence similarity to YdiQRST (Supplementary Table 5). This analysis suggests that ferredoxin and flavoproteins encoded by *ydiQRST*[20] are involved in the transfer of electrons to YdiO during the reduction of enoyl-CoA to acyl-CoA. Standard Gibbs energy calculations reveal that the engineered reversal of the β-oxidation cycle is thermodynamically feasible if ferredoxin is the source of reducing power for the conversion of enoyl-CoA to acyl-CoA (Supplementary Table 6). We then propose that the reduction of enoyl-CoA to acyl-CoA is mediated by YdiO-YdiQRST.

The engineered reversal of the β-oxidation cycle generates a diverse set of CoA thioester intermediates that can be converted to the corresponding alcohols and carboxylic acids (Fig. 1). To illustrate product synthesis from intermediates other than acyl-CoA, we used thioesterase I (TesA)[24] and thioesterase II (TesB)[25] as termination enzymes. Small amounts of 3-hydroxybutyric, 3-ketobutyric and trans-2-butenoic acids were produced when these thioesterases were overexpressed in strain RB02 (Fig. 2c). The level of these products was significantly increased by simultaneous overexpression of thioesterase and *yqeF*-encoded short-chain acyltransferase (Fig. 2c). Further increases in titre were realized upon deletion of *fadB* (~500 mg $l^{-1}$ 3-ketobutyric acid) and *ydiO* (~150 mg $l^{-1}$ and 200 mg $l^{-1}$ of 3-hydroxybutyric and trans-2-butenoic acids, respectively) (Fig. 2c).

The operation of multiple cycles of the engineered reversal of the β-oxidation pathway, and hence the synthesis of CoA-thioester intermediates (and products) of longer chain length ($C > 4$), can be facilitated by the overexpression of FadA, a 3-ketoacyl-CoA thiolase that is part of the β-oxidation complex (FadBA) and which possesses broad chain length specificity[26]. Overexpression of FadBA in conjunction with thioesterases (TesA, TesB, FadM[27] or YciA[28]) in strain RB03 (RB02 Δ*yqhD* Δ*fucO* Δ*fadD*) resulted in the accumulation of long-chain fatty acids in the extracellular medium (Fig. 3a). The *fadD* knockout in strain RB03 prevents re-utilization of the synthesized fatty acids. The choice of thioesterase allowed control over both length and functionality of the fatty-acid side chain. For example, C16 and C18 saturated fatty acids were the only products when FadM was overexpressed, whereas YciA and TesA overexpression supported the synthesis of 3-hydroxy (C14:3OH) and unsaturated (C18:1) fatty acids, respectively (Fig. 3a). When grown in a bioreactor using a higher initial concentration of glucose, strain RB03 [*fadBA*-*fadM*+] produced long-chain extracellular fatty acids at high titre (~7 g $l^{-1}$) and yield (0.28 g fatty acids per g total glucose consumed) using a mineral salts medium without rich nutrients (Fig. 3b). These results are better than reported previously using an engineered fatty-acid biosynthesis pathway[2]

**Table 1 | Characterization of the engineered reversal of the β-oxidation cycle during *n*-butanol synthesis**

| Strain | Enzyme activity (μmol per mg protein per min) ± s.d. | | | |
|---|---|---|---|---|
| | THL | HBD | CRT | BDH |
| (**a**) Activities of β-oxidation and butanol dehydrogenase enzymes in wild-type and engineered strains | | | | |
| MG1655 | ND | 0.002 ± 0.000 | ND | 0.014 ± 0.001 |
| RB02 (*fadR atoC(c) crp* ΔarcA Δpta ΔadhE ΔfrdA*) | 0.310 ± 0.079 | 0.304 ± 0.032 | 0.339 ± 0.049 | 0.004 ± 0.002 |
| RB02 [*yqeF*+ *fucO*+] | 0.498 ± 0.036 | 0.292 ± 0.013 | 0.334 ± 0.017 | 0.298 ± 0.020 |

| Strain | Butanol produced | | Glucose used | Cell growth |
|---|---|---|---|---|
| | Yield (g $g^{-1}$) | Concentration (g $l^{-1}$) | (g $l^{-1}$) | (g $l^{-1}$) |
| (**b**) Butanol synthesis, glucose utilization and cell growth in strain RB02 and its derivatives* | | | | |
| RB02 [*yqeF*+ *fucO*+] | 0.182 | 1.85 | 10.19 | 0.72 |
| Reaction (1): YqeF (predicted acyltransferase) | | | | |
| RB02 Δ*yqeF* [*fucO*+] | 0.019 | 0.10 | 5.19 | 0.32 |
| RB02 [*fucO*+] | 0.063 | 0.23 | 3.66 | 0.41 |
| RB02 Δ*yqeF* [*fucO*+ *yqeF*+] | 0.159 | 1.11 | 7.00 | 0.43 |
| Reactions (2) and (3): FadB (3-hydroxyacyl-CoA dehydrogenase and enoyl-CoA hydratase) | | | | |
| RB02 Δ*fadB* [*yqeF*+ *fucO*+] | ND | ND | 2.96 | 0.19 |
| RB02 Δ*fadB* [*yqeF*+ *fadB*+] | 0.157 | 0.92 | 5.86 | 0.52 |
| RB02 Δ*fadB* [*yqeF*+] | ND | ND | 1.84 | 0.11 |
| Reaction (4): YdiO (predicted acyl-CoA dehydrogenase/enoyl-CoA reductase) | | | | |
| RB02 Δ*ydiO* [*yqeF*+ *fucO*+] | 0.038 | 0.12 | 3.14 | 0.18 |
| RB02 Δ*ydiO* [*yqeF*+ *ydiO*+] | 0.163 | 0.97 | 5.94 | 0.57 |
| RB02 Δ*ydiO* [*yqeF*+] | 0.018 | 0.10 | 1.66 | 0.13 |
| Reaction (5): MhpF (aldehyde-forming acyl-CoA reductase) and FucO (*n*-butanol dehydrogenase) | | | | |
| RB02 Δ*fucO* [*yqeF*+] | 0.015 | 0.04 | 2.75 | 0.38 |
| RB02 Δ*fucO* [*yqeF*+ *fucO*+] | 0.151 | 0.85 | 5.64 | 0.51 |
| RB02 [*yqeF*+] | 0.088 | 0.35 | 4.00 | 0.68 |
| RB02 Δ*mhpF* [*yqeF*+ *fucO*+] | ND | ND | 0.66 | 0.07 |

BDH, butanol dehydrogenase; CRT, crotonase; HBD, hydroxy-acyl-CoA dehydrogenase; THL, thiolase.
ND, not detectable (minimum detectable activity was 0.001 μmol per mg protein per min; minimum detection level for butanol was 5.84 mg $l^{-1}$).
* Experiments were run for 24 h in shake flasks using glucose (2% w/v) minimal medium at 30 °C.



©2011 Macmillan Publishers Limited. All rights reserved

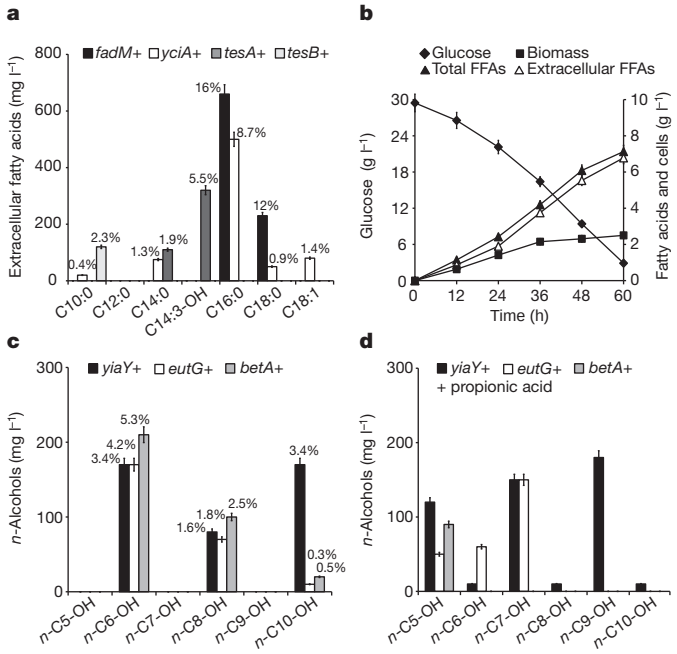



**Figure 3 | Synthesis of higher-chain (C > 4) carboxylic acids and *n*-alcohols through the engineered reversal of the β-oxidation cycle.** **a**, Accumulation of long-chain (C > 10) free fatty acids in the extracellular medium of strain RB03 [*fadBA*+] upon overexpression of different thioesterases (FadM, YciA, TesA, TesB). Product yield is shown above the bars (g free fatty acid per g total glucose consumed ×100). Experiments were run at 37 °C for 48 h in shake flasks using glucose (2% w/v) minimal medium. **b**, Kinetics of fatty-acid synthesis by strain RB03 [*fadBA-fadM*+]. Cells were cultivated in fermentors using glucose (3% w/v) minimal medium. The dissolved oxygen was controlled at 2% of saturation, 37 °C, pH 7. FFAs, free fatty acids. **c**, Synthesis of *n*-alcohols in strain RB03 [*fadBA*+] upon overexpression of alcohol dehydrogenases (YiaY, BetA and EutG). Product yield is shown above the bars (g *n*-alcohol per g total glucose consumed ×100). Experiments were run at 37 °C for 48 h in shake flasks using glucose (2% w/v) minimal medium. **d**, Effect of alcohol dehydrogenase overexpression (YiaY, BetA and EutG) on the chain-length distribution of *n*-alcohols synthesized by strain RB03 [*fadBA*+] in the presence of 0.5 $g\,l^{-1}$ propionate. Experiments were conducted as described in **c**. Error bars represent the standard deviation of three replicates.

(Supplementary Table 4). No production of extracellular free fatty acids was observed upon overexpression of FadM in strain MG1655 Δ*adhE* Δ*pta* Δ*frd*A Δ*fadD* (Supplementary Table 2C), demonstrating the requirement of an active reversal of the β-oxidation cycle. Measurements of total free fatty acids (that is, extracellular plus intracellular) in strain RB03 [*fadBA-fadM*+] and the corresponding controls showed that the engineered reversal of the β-oxidation cycle contributed to the synthesis of 90–95% of the total free fatty acids (Supplementary Table 2C).

The synthesis of longer-chain (C > 4) *n*-alcohols was also demonstrated by overexpressing the appropriate termination enzymes (Fig. 3c). We identified native enzymes that could serve as potential surrogates for the aldehyde-forming acyl-CoA reductases and alcohol dehydrogenases present in organisms that synthesize higher-chain linear *n*-alcohols (Supplementary Table 3). The product titre (0.33 $g\,l^{-1}$) and yield (8.3% w/w) achieved upon overexpression of YiaY were higher than previously reported[2,29] (Supplementary Table 4). Synthesis of odd-chain *n*-alcohols was demonstrated by supplementing the medium with propionate as the precursor of propionyl-CoA (R = $CH_3$ in Fig. 1). A clear shift in the distribution of *n*-alcohols was observed: odd-chain alcohols 1-pentanol, 1-heptanol and 1-nonanol appeared as fermentation products and the synthesis of even-chain alcohols significantly decreased (Fig. 3d).

The functional reversal of the β-oxidation cycle engineered in this work represents a novel and efficient platform for the synthesis of advanced fuels and chemicals. Its superior nature is illustrated in the following balanced equation for the synthesis of *n*-alcohols from glucose: $n/4\ C_6H_{12}O_6 \rightarrow C_nH_{n+2}O + n/2\ CO_2 + (n/2 - 1)\ H_2O + n/2$ ATP, with *n* being the chain length of the *n*-alcohol (Fig. 1). As can be seen in this balance, the engineered pathway has the potential to achieve the maximum theoretical yield of *n*-alcohols on glucose (66.7%, C-mole basis) and generates 1 ATP per each two-carbon unit incorporated into the *n*-alcohol molecule. This ATP yield is equivalent to that of efficient homo-fermentative pathways found in nature such as ethanol[30] and lactic acid[19] fermentations. The high efficiency of the engineered reversal of the β-oxidation cycle is possible because it directly uses acetyl-CoA as the donor of two-carbon units during chain elongation (as opposed to first requiring ATP-dependent activation to malonyl-CoA) and it functions with acyl-CoA intermediates, which are the precursors of *n*-alcohols and other important products (Fig. 1).

The synthesis of *n*-alcohols through alternative metabolic routes, such as the fatty-acid biosynthesis[2] and keto-acid[29] pathways, is less efficient. The use of the fatty acid biosynthesis pathway results in the net consumption of 1 ATP per molecule of *n*-alcohol synthesized (Supplementary Fig. 1). This inefficiency is due to the consumption of ATP in the synthesis of malonyl-ACP, the donor of two-carbon units for chain elongation, and the use of acyl-ACP intermediates, which need to be converted to free acids and acylated (ATP-consuming step) before their reduction to alcohols can be achieved (Supplementary Fig. 1). The keto-acid pathway[29] is also less efficient than the reversal of the β-oxidation cycle: for example, the maximum theoretical yield of *n*-hexanol, the highest-chain linear *n*-alcohol reported with the keto-acid pathway, is only 50% C-mole (2 glucose → *n*-hexanol + ATP + 2[H] + 6 $CO_2$).

## METHODS SUMMARY

**Strains and plasmids.** Wild-type K12 *Escherichia coli* strain MG1655 was used as the host to implement metabolic engineering strategies. Genetic manipulations, including construction of plasmids, are described in the Methods. All resulting strains and plasmids are listed in Supplementary Table 7.

**Culture medium, cultivation conditions and analytical methods.** The mineral salts medium described in Methods was used. Ampicillin (100 μg $ml^{-1}$), kanamycin (50 μg $ml^{-1}$), chloramphenicol (34 μg $ml^{-1}$), anhydrotetracycline (100 ng $ml^{-1}$) and isopropyl β-D-1-thiogalactopyranoside (IPTG; 0.1 mM) were included when appropriate. Several colonies were used to inoculate 30 ml of medium in 50 ml baffled flasks and incubated at 37 °C in a rotary shaker (200 r.p.m.) for 24 h. An appropriate amount of pellet from this pre-culture was used to inoculate 30 ml of fresh medium in 50 ml baffled flasks to a target initial optical density at 550 nm of 0.05. These cultures were then incubated at 30 °C in a rotary shaker (200 r.p.m.) for 24 h, unless otherwise stated. The same inoculation procedure was used for additional fermentations conducted in a SixFors multi-fermentation system (Infors HT) with an air flowrate of 1.5 $N\,l^{-1}\,h^{-1}$ and control of dissolved oxygen, pH, and temperature at specified values. Metabolite identification and quantification by nuclear magnetic resonance (NMR), gas chromatography-mass spectroscopy (GC-MS), gas chromatography-flame ionization detection (GC-FID), and high-performance liquid chromatography (HPLC) are described in Methods. Product yields were calculated as grams of product synthesized per total grams of glucose consumed. All values reported as different are significant with a *P*-value lower than 0.05.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 21 January; accepted 23 June 2011.**
**Published online 10 August 2011.**

1. Haapalainen, A. M., Merilainen, G. & Wierenga, R. K. The thiolase superfamily: condensing enzymes with diverse reaction specificities. *Trends Biochem. Sci.* **31,** 64–71 (2006).
2. Steen, E. J. *et al.* Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463,** 559–562 (2010).
3. Handke, P., Lynch, S. A. & Gill, R. T. Application and engineering of fatty acid biosynthesis in *Escherichia coli* for advanced fuels and chemicals. *Metab. Eng.* **13,** 28–37 (2010).
4. Lennen, R. M., Braden, D. J., West, R. M., Dumesic, J. A. & Pfleger, B. F. A process for microbial hydrocarbon synthesis: Overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes. *Biotechnol. Bioeng.* **106,** 193–202 (2010).



©2011 Macmillan Publishers Limited. All rights reserved

5. Clark, D. P. & Cronan, J. E. in Escherichia coli *and* Salmonella*: Cellular and Molecular Biology* (ed. Neidhart, F. C.) Ch. 3.4.4 343–357 (American Society for Microbiology, 2005).
6. Poirier, Y., Antonenkov, V. D., Glumoff, T. & Hiltunen, J. K. Peroxisomal β-oxidation—A metabolic pathway with multiple functions. *Biochim. Biophys. Acta* **1763,** 1413–1426 (2006).
7. Shen, C. R. & Liao, J. C. Metabolic engineering of *Escherichia coli* for 1-butanol and 1-propanol production via the keto-acid pathways. *Metab. Eng.* **10,** 312–320 (2008).
8. Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451,** 86–89 (2008).
9. Bond-Watts, B. B., Bellerose, R. J. & Chang, M. C. Y. Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nature Chem. Biol.* **7,** 222–227 (2011).
10. Dellomonaco, C., Rivera, C., Campbell, P. & Gonzalez, R. Engineered respiro-fermentative metabolism for the production of biofuels and biochemicals from fatty acid-rich feedstocks. *Appl. Environ. Microbiol.* **76,** 5067–5078 (2010).
11. Deutscher, J. The mechanisms of carbon catabolite repression in bacteria. *Curr. Opin. Microbiol.* **11,** 87–93 (2008).
12. Eppler, T. & Boos, W. Glycerol-3-phosphate-mediated repression of *malT* in *Escherichia coli* does not require metabolism, depends on enzyme IIA(Glc) and is mediated by cAMP levels. *Mol. Microbiol.* **33,** 1221–1231 (1999).
13. Cho, B. K., Knight, E. M. & Palsson, B. O. Transcriptional regulation of the *fad* regulon genes of *Escherichia coli* by ArcA. *Microbiology-Sgm* **152,** 2207–2219 (2006).
14. Jarboe, L. YqhD: a broad-substrate range aldehyde reductase with various applications in production of biorenewable fuels and chemicals. *Appl. Microbiol. Biotechnol.* **89,** 249–257 (2010).
15. Feigenbaum, J. & Schulz, H. Thiolases of *Escherichia coli*: Purification and chain length specificities. *J. Bacteriol.* **122,** 407–411 (1975).
16. Jenkins, L. S. & Nunn, W. D. Genetic and molecular characterization of the genes involved in short-chain fatty acid degradation in *Escherichia coli*: the *ato* system. *J. Bacteriol.* **169,** 42–52 (1987).
17. Lee, S. Y. *et al.* Fermentative butanol production by clostridia. *Biotechnol. Bioeng.* **101,** 209–228 (2008).
18. Gonzalez, R., Tao, H., Shanmugam, K. T., York, S. W. & Ingram, L. O. Global gene expression differences associated with changes in glycolytic flux and growth rate in *Escherichia coli* during the fermentation of glucose and xylose. *Biotechnol. Prog.* **18,** 6–20 (2002).
19. Papagianni, M., Avramidis, N. & Filiousis, G. Glycolysis and the regulation of glucose transport in *Lactococcus lactis* spp. *lactis* in batch and fed-batch culture. *Microb. Cell Fact.* **6,** doi:10.1186/1475-2859-6-16 (2007).
20. Campbell, J. W., Morgan-Kiss, R. M. & Cronan, J. E. A new *Escherichia coli* metabolic competency: growth on fatty acids by a novel anaerobic β-oxidation pathway. *Mol. Microbiol.* **47,** 793–805 (2003).
21. Eichler, K., Bourgis, F., Buchet, A., Kleber, H. P. & Mandrandberthelot, M. A. Molecular characterization of the *cai* operon necessary for carnitine metabolism in *Escherichia coli*. *Mol. Microbiol.* **13,** 775–786 (1994).
22. Walt, A. & Kahn, M. L. The *fixA* and *fixB* genes are necessary for anaerobic carnitine reduction in *Escherichia coli*. *J. Bacteriol.* **184,** 4044–4047 (2002).
23. Eichler, K., Buchet, A., Bourgis, F., Kleber, H. P. & Mandrandberthelot, M. A. The *fix Escherichia coli* region contains four genes related to carnitine metabolism. *J. Basic Microbiol.* **35,** 217–227 (1995).
24. Cho, H. S. & Cronan, J. E. *Escherichia coli* thioesterase I, molecular cloning and sequencing of the structural gene and identification as a periplasmic enzyme. *J. Biol. Chem.* **268,** 9238–9245 (1993).
25. Nie, L., Ren, Y., Janakiraman, A., Smith, S. & Schulz, H. A novel paradigm of fatty acid β-oxidation exemplified by the thioesterase-dependent partial degradation of conjugated linoleic acid that fully supports growth of *Escherichia coli*. *Biochemistry* **47,** 9618–9626 (2008).
26. Yang, S. Y., Yang, X. Y. H., Healylouie, G., Schulz, H. & Elzinga, M. Primary structure of 3-ketoacyl-coenzyme A thiolase from *Escherichia coli* and the structural organization of the *fadAB* operon. *J. Biol. Chem.* **265,** 10424–10429 (1990).
27. Feng, Y. & Cronan, J. E. A new member of the *Escherichia coli fad* regulon: Transcriptional regulation of *fadM* (*ybaW*). *J. Bacteriol.* **191,** 6320–6328 (2009).
28. Zhuang, Z. H. *et al.* Divergence of function in the hot dog fold enzyme superfamily: The bacterial thioesterase YciA. *Biochemistry* **47,** 2789–2796 (2008).
29. Zhang, K. C., Sawaya, M. R., Eisenberg, D. S. & Liao, J. C. Expanding metabolism for biosynthesis of nonnatural alcohols. *Proc. Natl Acad. Sci. USA* **105,** 20653–20658 (2008).
30. Osman, Y. A., Conway, T., Bonetti, S. J. & Ingram, L. O. Glycolytic flux in *Zymomonas mobilis*: enzyme and metabolite levels during batch fermentation. *J. Bacteriol.* **169,** 3726–3736 (1987).


**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank B. Erni and H. Mori for providing research materials, S. Moran and J. F. Fallas Valverde for assistance with NMR techniques, and D. A. Castillo-Rivera, B. Wilson, S. P. T. Matsuda, M. Li and K.-Y. San for assistance with GC-MS techniques. R.G. thanks N. E. Gonzalez, B. C. Gutierrez and M. D. Diaz for their continued support.

**Author Contributions** R.G. conceived the work. C.D. and R.G. designed the experiments. C.D. conducted all strain characterization experiments. C.D. performed the *in silico* analyses. C.D., J.M.C. and E.N.M. constructed the strains. J.M.C. performed enzyme assays and the thermodynamic analysis. R.G. and C.D. drafted the manuscript. All authors read, edited and approved the final manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.G. (Ramon.Gonzalez@rice.edu).



©2011 **Macmillan Publishers Limited. All rights reserved**

## METHODS

**Strains and plasmids.** Gene knockouts were introduced in MG1655 and its derivatives by P1 phage transduction as previously described[31]. The native cAMP receptor protein (CRP) was replaced with a cAMP-independent CRP (CRP*)[32] by introducing the *crp** gene into *crp* deletion mutants via P1 phage transduction using a lysate from strain ET25 (ref. 13), followed by selection on (1% w/v) glycerol minimal media plates with tetracycline. Construction of plasmids (medium-copy pTrcHis2A, Invitrogen, and low-copy pZSblank (ref. 31)) is described in the 'Genetic methods' section. All resulting strains and plasmids are listed in Supplementary Table 7.

**Reagents.** Chemicals were obtained from Fisher Scientific and Sigma-Aldrich.

**Culture medium.** The minimal medium designed previously[33] with $Na_2HPO_4$ in place of $K_2HPO_4$ and supplemented with 20 g $l^{-1}$ glucose, 40 g $l^{-1}$ calcium carbonate, 100 μM $FeSO_4$, 5 mM calcium pantothenate, 3.96 mM $Na_2HPO_4$, 5 mM $(NH_4)_2SO_4$ and 30 mM $NH_4Cl$ was used. Fermentations conducted in the SixFors multi-fermentation system also included 1 mM betaine. Ampicillin (100 μg $ml^{-1}$), kanamycin (50 μg $ml^{-1}$), chloramphenicol (34 μg $ml^{-1}$), anhydrotetracycline (100 ng $ml^{-1}$) and isopropyl β-D-1-thiogalactopyranoside (IPTG; 0.1 mM) were included when appropriate. Several colonies were used to inoculate 30 ml of medium in 50 ml baffled flasks and incubated at 37 °C in a rotary shaker (200 r.p.m.) for 24 h. An appropriate amount of pellet from this pre-culture was used to inoculate 30 ml of fresh medium in 50 ml baffled flasks to a target initial optical density at 550 nm of 0.05. These cultures were then incubated at 30 °C in a rotary shaker (200 r.p.m.) for 24 h, unless otherwise stated.

**Genetic methods.** Standard recombinant DNA procedures were used for gene cloning, plasmid isolation and electroporation[34]. Manufacturer protocols and standard methods were followed for DNA purification (Qiagen), restriction endonuclease digestion (New England Biolabs) and DNA amplification (Stratagene and Invitrogen). For plasmid construction, genes were amplified from MG1655 genomic DNA using primers designed to create 15 bp of homology on each end of the gene insert for subsequent recombination into the desired plasmid. Plasmids were linearized using restriction endonuclease digestion, then recombined with the appropriate gene(s) using an In-Fusion Dry-Down PCR Cloning Kit (Clontech) and subsequently used to transform chemically competent Fusion Blue cells (Clontech). Transformants that grew on LB plates containing the appropriate antibiotic were struck for isolation, and then subjected to preliminary screening by PCR. Colonies passing preliminary inspection were then individually grown for plasmid purification. Purified plasmids were confirmed to have the appropriate insert both by PCR as well as restriction endonuclease digest verification. Plasmids in each case include the plasmid-based promoter, a ribosomal binding site for each gene, MG1655 gene(s), and a plasmid-based terminator. Resulting plasmids are listed in Supplementary Table 7.

**Sequencing of regulators.** Chromosomal DNA from strain RB02 served as the template for PCR using Phusion High-Fidelity DNA polymerase (Finnzymes). PCR products, containing the gene(s) of interest as well as approximately 500 bp upstream of each gene, were purified using a QIAquick PCR purification kit (Qiagen) and sent to Lone Star Labs for sequencing. Sequencing results were analysed using Sequence Scanner v1.0 (Applied Biosystems) and Serial Cloner v2.1 (Serial Basics, http://serialbasics.free.fr/Home/Home.html).

**Metabolite identification.** The identity of metabolic products was determined through one-dimensional proton nuclear magnetic resonance (NMR) spectroscopy. 60 μl of $D_2O$ and 1 μl of 600 mM NMR internal standard TSP [3-(trimethylsilyl) propionic acid-D4, sodium salt] were added to 540 μl of the sample (culture supernatant). The resulting solution was then transferred to a 5-mm NMR tube, and one-dimensional proton NMR spectroscopy was performed at 25 °C in a Varian 500-MHz Inova spectrometer equipped with a Penta probe (Varian) using the following parameters: 8,000-Hz sweep width, 2.8-s acquisition time, 256 acquisitions, 6.3 μs pulse width, 1.2-s pulse repetition delay, and pre-saturation for 2 s. The resulting spectrum was analysed using FELIX 2001 software (Accelrys Software). Peaks were identified by their chemical shifts and J-coupling values, which were obtained in separate experiments in which samples were spiked with metabolite standards (2 mM final concentration).

Identification of *n*-alcohols was conducted through GC-MS following a modification of the method reported previously[35]. The analysis was performed on an Agilent 6890 GC/5973 MS (Agilent Technologies) instrument with a HP-5ms capillary column (30 m × 0.25 mm × 0.25 μm). 1 ml of supernatant of culture broth was extracted with 500 μl of GC standard grade *n*-hexane (Fluka) and 1-heptanol (Sigma-Aldrich) was used as internal standard. 0.5 μl of the extracted sample was injected using a 20:1 split at 250 °C. The oven temperature was initially held at 75 °C for 2 min and then raised with a gradient of 5 °C per min to 280 °C and held for 2 min. Helium (Matheson Tri-Gas) was used as the carrier gas with a 14-lb $in^{-2}$ inlet pressure. The injector and detector were maintained at 255 °C.

Identification of fatty acids was performed on a Shimadzu Auto-System GC 2010 (Shimadzu) equipped with a DB-5MS capillary column (30 m × 0.25 mm × 0.25 μm) and coupled with a QP-2010 mass detector. The following method was used: an initial temperature of 50 °C was held for 2 min and then ramped to 220 °C at 4 °C per min and held for 10 min (ref. 2). Extraction and derivatization procedures are described in the 'Metabolite quantification' section.

**Metabolite quantification.** The quantification of glucose, organic acids, ethanol and butanol was conducted by high-performance liquid chromatography (HPLC). Samples (culture supernatant) were analysed with ion-exclusion HPLC using a Shimadzu Prominence SIL 20 system (Shimadzu Scientific Instruments) equipped with an HPX-87H organic acid column (Bio-Rad) with operating conditions to optimize peak separation (0.3 ml $min^{-1}$ flowrate, 30 mM $H_2SO_4$ mobile phase, column temperature 42 °C)[36].

Quantification of longer chain (C > 4) *n*-alcohols was conducted through gas chromatography (GC) in a Varian CP-3800 gas chromatograph (Varian Associates) equipped with a flame ionization detector (GC-FID). Sample extraction procedure was as described above. The separation of alcohol compounds was carried out using a FactorFour VF-5ht column (15 m, 0.32 μm internal diameter, 0.10 μm film thickness; Varian Associates). The oven temperature was initially held at 40 °C for 1 min and then raised with a gradient of 30 °C $min^{-1}$ to 130 °C and held for 4 min. The temperature was then raised with a gradient of 15 °C $min^{-1}$ to 230 °C and held for 4 min. Helium (1 ml $min^{-1}$, Matheson Tri-Gas) was used as the carrier gas. The injector and detector were maintained at 250 °C. A 0.5-μl sample was injected in splitless injection mode.

For total free fatty acid quantification, 5 ml of culture was placed in a 15-ml Corning CentriStar centrifuge tube (Corning Inc.) in an ice bath and sonicated using a Branson Sonifier 450 (Emerson Industrial Automation) for 5 min at a duty cycle of 30% and a power setting of 1. The tubes were then centrifuged at 12,000*g* for 5 min and the appropriate volume of supernatant recovered. For extracellular free fatty-acid analysis, the supernatant from non-sonicated culture samples was used instead. 2 ml of supernatant were acidified with 200 μl of acetic acid, supplemented with 100 mg of octanoic acid as an internal standard, and extracted with 2 ml of *n*-hexane-chlorophorm (4:1, v/v)[37]. The organic layer was nitrogen evaporated to near dryness, re-dissolved in 1 ml of a mixture of cholophorm:methanol: hydrochloric acid (10:1:1, vol/vol/vol)[10] and incubated in a sealed vial at 90 °C for 1 h. Fatty-acid methyl esters were extracted with 1 ml of *n*-hexane-chlorophorm (4:1, v/v) after addition of 500 μl of 0.9% (wt/vol) NaCl in water. Gas-liquid chromatography was performed with 1 μl of the *n*-hexane:chlorophorm solution on a Varian FactorFour VF-5ht capillary column (Varian Associates, 15 m length, 0.32 μm diameter and 0.1 μm film thickness) according to the following method: 50 °C held for 1 min, 30 °C $min^{-1}$ to 160 °C, 15 °C $min^{-1}$ to 200 °C, 200 °C held for 1.5 min, 10 °C $min^{-1}$ to 225 °C, and 225 °C held for 15 min (ref. 10).

**Enzyme assays.** For measurement of enzymatic activities, cells from 24-h shake flask cultures were washed twice with 9 g $l^{-1}$ sodium chloride under anaerobic conditions and stored at −80 °C until use. Cell extracts for all assays were prepared as follows under anaerobic conditions: 40 units of $OD_{550\,nm}$ were re-suspended in 1 ml of 100 mM Tris-HCl buffer (pH 7.0) with 1 mM DTT. After cellular disruption using a Disruptor Genie (Scientific Industries), cellular debris was removed by centrifugation (13,000*g*, 4 °C, 10 min) and the supernatant used as cell extract. Absorbance changes for all assays were monitored in a Biomate 5 spectrophotometer (Thermo Scientific). The linearity of reactions (protein concentration and time) was established for all assays and the non-enzymatic rates were subtracted from the observed initial reaction rates. Enzymatic activities are reported as μmol of substrate per min per mg of cell protein and represent averages for at least three cell preparations. Protein concentration was measured using the Bradford Assay Reagent (Thermo Scientific) with BSA as a standard.

Acetyl-CoA acetyltransferase (THL) activity was determined using acetoacetyl-CoA and CoA as substrates, and the decrease in acetoacetyl-CoA concentration was measured at 303 nm (ref. 38). β-Hydroxybutyryl-CoA dehydrogenase activity was measured at 340 nm by monitoring the decrease in NADH concentration resulting from β-hydroxybutyryl-CoA formation from acetoacetyl-CoA (ref. 39). Crotonase activity was measured by monitoring the decrease in crotonyl-CoA concentration at 263 nm, which results from β-hydroxybutyryl-CoA formation from crotonyl-CoA (ref. 39). To measure butanol dehydrogenase activity, the decrease in NAD(P)H concentration resulting from butanol formation from butyraldehyde is monitored at 340 nm under anaerobic conditions at 30 °C (ref. 40).


31. Yazdani, S. S. & Gonzalez, R. Engineering *Escherichia coli* for the efficient conversion of glycerol to ethanol and co-products. *Metab. Eng.* **10,** 340–351 (2008).
32. Cirino, P. C., Chin, J. W. & Ingram, L. O. Engineering *Escherichia coli* for xylitol production from glucose-xylose mixtures. *Biotechnol. Bioeng.* **95,** 1167–1176 (2006).
33. Neidhart, F. C., Bloch, P. L. & Smith, D. F. Culture media for enterobacteria. *J. Bacteriol.* **119,** 736–747 (1974).


©2011 Macmillan Publishers Limited. All rights reserved

34. Sambrook, J. & Russell, D. W. *Molecular Cloning: a Laboratory Manual* (Cold Spring Harbor Laboratory, 2001).
35. Atsumi, S. *et al.* Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metab. Eng.* **10,** 305–311 (2008).
36. Dharmadi, Y. & Gonzalez, R. A better global resolution function and a novel iterative stochastic search method for optimization of high-performance liquid chromatographic separation. *J. Chromatogr. A* **1070,** 89–101 (2005).
37. Lalman, J. A. & Bagley, D. M. Extracting long-chain fatty acids from a fermentation medium. *J. Am. Oil Chem. Soc.* **81,** 105–110 (2004).
38. Wiesenborn, D. P., Rudolph, F. B. & Papoutsakis, E. T. Thiolase from *Clostridium acetobutylicum* ATCC824 and its role in the synthesis of acids and solvents. *Appl. Environ. Microbiol.* **54,** 2717–2722 (1988).
39. Hartmanis, M. G. N. & Gatenbeck, S. Intermediary metabolism in *Clostridium acetobutylicum*—levels of enzymes involved in the formation of acetate and butyrate. *Appl. Environ. Microbiol.* **47,** 1277–1283 (1984).
40. Durre, P., Kuhn, A., Gottwald, M. & Gottschalk, G. Enzymatic investigations on butanol dehydrogenase and butyraldehyde dehydrogenase in extracts of *Clostridium acetobutylicum*. *Appl. Microbiol. Biotechnol.* **26,** 268–272 (1987).

©2011 **Macmillan Publishers Limited. All rights reserved**

CORRIGENDUM

doi:10.1038/nature10338

# Structure and mechanism of human DNA polymerase η

Christian Biertümpfel, Ye Zhao, Yuji Kondo, Santiago Ramón-Maiques, Mark Gregory, Jae Young Lee, Chikahide Masutani, Alan R. Lehmann, Fumio Hanaoka & Wei Yang

*Nature* **465,** 1044–1048 (2010)

It was brought to our attention by D. Guillaume (Chimie Thérapeutique, CNRS) that deoxyribose C4′ of the cyclobutane pyrimidine dimer (CPD) photoproduct in the TT2 structure (Protein Data Bank accession number PDB 3MR4) has an inverted configuration. After further refinement, we obtained the correct sugar configuration (see Fig. 1 below). A new accession number (PDB 3SI8) has been given to the corrected structure.

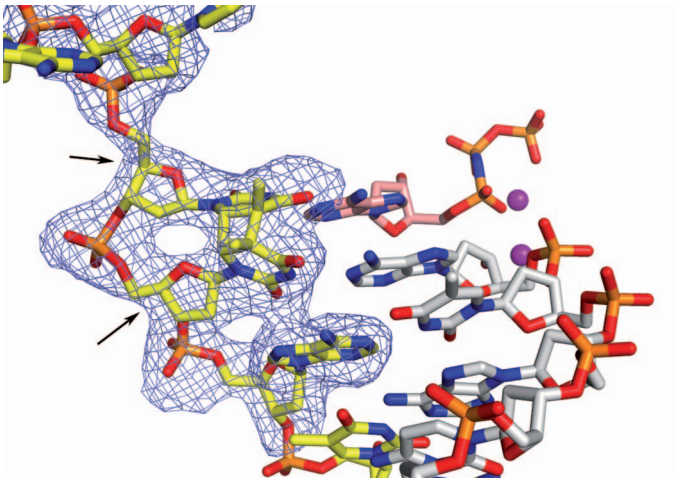


**Figure 1** | The template (yellow), primer (grey) and incoming nucleotide (pink) are shown as sticks. The omit map of the thymine dimer and surrounding nucleotides is shown as a blue mesh. Arrowheads indicate the C4′ of the crosslinked bases.



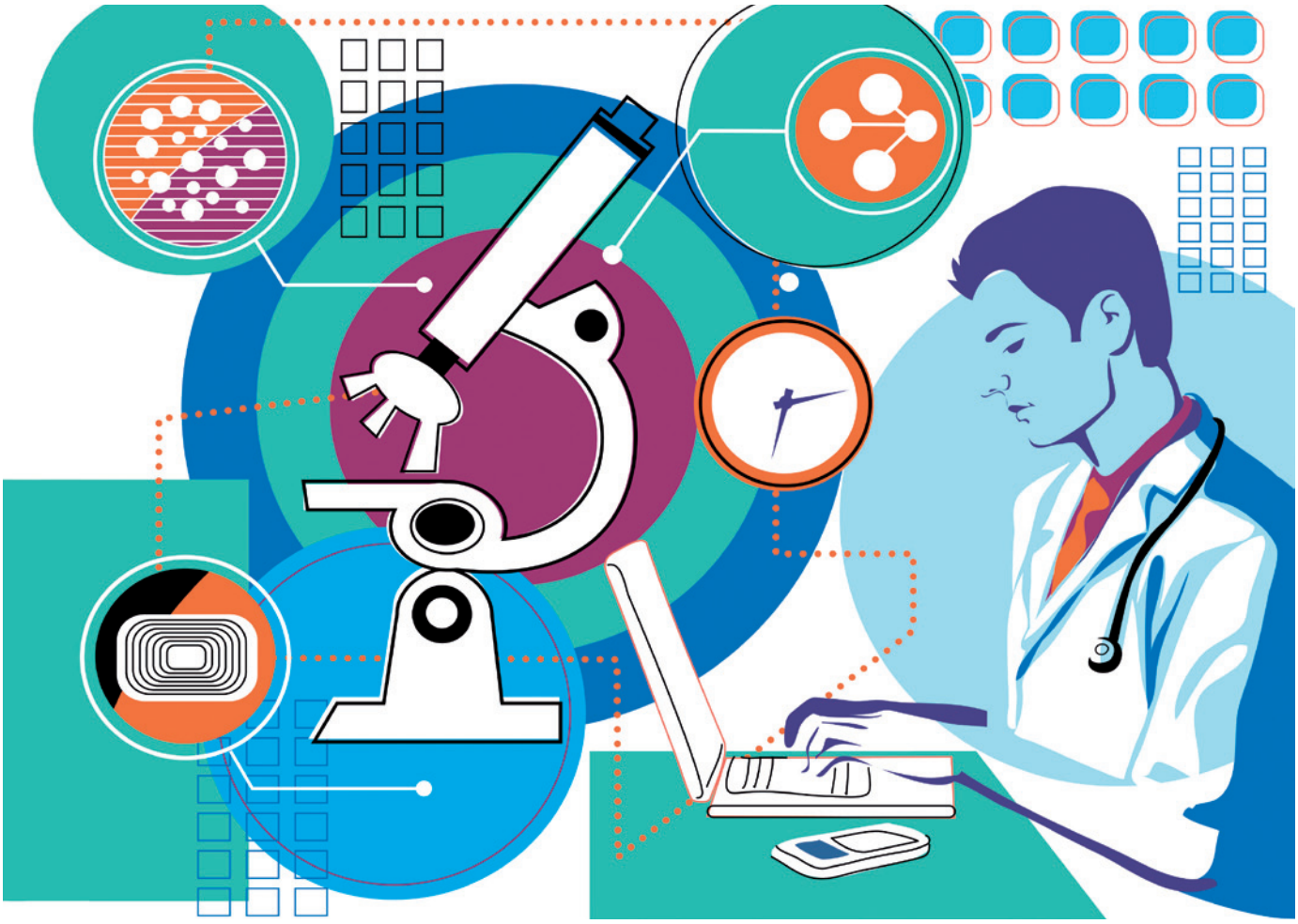
©2011 Macmillan Publishers Limited. All rights reserved

# CAREERS

**UNITED STATES** Fellowship supports international graduate students. **p.363**

**EDUCATION** Number of physics PhDs on the rise in the United States. **p.363**

**NATUREJOBS** For the latest career listings and advice **www.naturejobs.com**



CORBIS

CLINICAL SCIENCE

# Research and repair

*Maintaining a research focus is among the challenges facing aspiring physician–scientists.*

BY JEFFREY PERKEL

For every day that Thomas Cappola spends tending to patients, he spends four in his lab, trying to understand at the genetic level why those patients are ill, and how they can best be treated. To do so, Cappola, a cardiologist at the University of Pennsylvania School of Medicine in Philadelphia, relies not only on his extensive medical training, but also on a postdoctoral fellowship in applied genomics.

Cappola is a physician-scientist — a scientifically trained clinician who works in both patient care and research, in this case applying genomics technology, bioinformatics and statistics to his heart-failure patients' blood and DNA samples. Cappola likes to say that he has a "crazy set of skills", which he enjoys using. "I see patients one day a week, and the rest of the time I get to think about their problems and how to solve them," he says.

In a sense, physician-scientists are like any other scientist competing for the jobs, funding and personnel required to advance their research. Yet theirs is a profession with its own hurdles and opportunities, and one to which funding agencies are paying special attention, especially with taxpayers demanding proof that federal grants are being put to good use.

The career path isn't easy. It requires a diverse set of research and people skills, and a willingness to endure long hours, intensive schooling and, frequently, multiple degree programmes. But those who secure degrees and physician-scientist posts may find it easier to get jobs than do most people with PhDs in the life sciences. ▶



© 2011 Macmillan Publishers Limited. All rights reserved

Physician-scientists can also boast a satisfying vocation that combines curiosity-driven research with a patient-focused approach. Aspirants should beware, however: in some quarters, an emphasis on translational research, stemming in part from economic realities, has forced physician-scientists to spend more time in the clinic than they might have in the past.

## LONG AND EXPENSIVE

No hard numbers on career prospects exist, but anecdotal evidence suggests that physician-scientists compete well with their PhD-holding peers. In part, that is because the physician-scientist pool, at least in the United States, has been contracting and ageing for years, says Andrew Schafer, chairman of the department of medicine at Weill Cornell Medical College in New York and editor of *The Vanishing Physician–Scientist?* (ILR, 2010).

"There's no question that there is a critical situation," says Schafer. "The disappearance of physician-scientists threatens the very essence of what we call translational research, because if not for them, who would be left to translate anything?" As a result, he says, job prospects in general are good. "Virtually every major academic medical centre wants physician-scientists," says Schafer.



***"Virtually every major academic medical centre wants physician-scientists."***
Andrew Schafer

However, there are myriad barriers to establishing oneself as a physician-scientist. First, it can require a long stint at university. For example, the bachelor of medicine (MB)/PhD programme at University College London (UCL) follows a 0.5–3–2 format — after an undergraduate degree that includes preclinical coursework, the student undertakes half a year of postgraduate clinical training, then 2.5–3.5 years of research and another two years of clinical work. By contrast, US combined MD–PhD programmes, initiated after completion of an undergraduate non-clinical programme, typically take 8 years in a 2–4–2 format: two years of preclinical coursework, four of research and two of clinical training.

For those who do not receive funding as part of their MD–PhD programmes (which, at least in the United States, are tuition-free and include a stipend), university can be expensive. US physicians who want to focus on research must often deal with a mountain of debt as they pursue further studies or a research postdoc. According to a 2010 survey[1] by the Association of American Medical Colleges (AAMC) in Washington DC, the average indebted medical student owes nearly US$160,000 at graduation. Many are therefore eager to get into a practice and start paying off their loans. "The vast majority of MDs have so much debt they cannot afford to do research unless they are either independently wealthy or are very, very motivated," says Ajit Varki, a biomedical scientist at the University of California, San Diego, and former associate dean of the university's physician-scientist training programme.

The US National Institutes of Health (NIH) has a loan-repayment programme (LRP) that helps to defray some costs for physicians who decide to move into research, paying back up to $35,000 of debt per year on the recipient's behalf. Cappola used an LRP, but it didn't cover the entire balance of his debt. "I am debt-free now," he says, "but it took time and not just the LRP. It took a lot of work."



UNIV. PENNSYLVANIA

***"I get to see patients one day a week and the rest of the time I get to think about their problems and how to solve them."***
Thomas Cappola

UCL's MB/PhD programme is tuition-free for the research period, but there are charges for the clinical training. Still, says Gordon Stewart, programme tutor at UCL, the cost is much lower than that of straight US medical degrees; UK medical students typically graduate with £15,000–40,000 (US$24,000–65,000) of debt, he says. "The loans are not nearly as big as they have to take out in the United States, but the salaries they get at the end of the day are nothing like as big as they get in the United States, either," he says.

WEILL CORNELL MED. COLL.

Once they get past university, physician-scientists must survive not only the transition from postdoc to faculty member, but also the initial phases of their first posts, as they try to establish self-perpetuating research programmes. This entails an ongoing battle over the pressure to work in the clinic. Most physician-scientists have to cover their salary costs — the patient hours that must be billed to match the cost of their salary. Granting agencies have attempted to support physician-scientists through several different funding mechanisms,

## GRANT OPPORTUNITIES

### *Mechanisms of support*

There is plenty of interest in fostering physician-scientists, and many funding possibilities. In the United States, for example, the National Institutes of Health (NIH) offers awards known as K grants, some of which fund budding physician-scientists. Awarded separately by each NIH institute, the schemes support licensed clinicians, dentists or veterinarians in basic (the K08 grant) or clinical (K23) mentored research projects, requiring the recipient to spend at least 75% of their time on research, during the difficult period between completion of their research training as a postdoc and the acquisition of principal-investigator grants.

For earlier-stage training, the Howard Hughes Medical Institute (HHMI), based in Chevy Chase, Maryland, offers one-year fellowships at either a host university or the NIH, called the Medical Research Fellows Program and the HHMI-NIH Research Scholars Program, respectively, to give medical, dental and veterinary students a chance to try their hand at research.

In the United Kingdom, Intermediate and Senior Clinical Fellowships from the London-based Wellcome Trust support physician-scientists during PhD education, postdoctoral training and as principal investigators. Cancer Research UK, also based in London, offers postdoc and senior fellowships to protect research time. The charity awarded five four-year Clinician Scientist Fellowships in 2011, each supporting the researcher's salary, a research assistant and supplies; the Senior Cancer Fellowship supports salary, two postdocs, a technician, a graduate student and consumables for six years.

Kathleen Cook Suozzi, a physician-scientist-in-training, says that her HHMI fellowship experience helped her to see how practicing physician-scientists manage their double lives, in part through conferences organized by the institute. "The spectrum upon which people split their time is pretty vast," says Cook Suozzi, a fourth-year medical student at New York University School of Medicine, who spent a year in a lab at Rockefeller University in New York. "I met physicians who spend 99% of their time in the lab, and I met surgical oncologists with a lab who are in the operating room 80% of the time. So we could see that you can make it work in a variety of ways and make it fit your needs." J.P.


© 2011 Macmillan Publishers Limited. All rights reserved

available to early-career scientists as well as seasoned practitioners (see 'Mechanisms of support'). But if they fail to win such grants, physician-scientists can be required to fall back on their medical training to see more patients, thereby diminishing their time in the lab. Down-on-their-luck physician-scientists can thus become progressively less competitive in research.

### TOUGH TRANSITION

Supporting both research and clinical priorities is not easy, as Melina Kibbe, a vascular surgeon at Northwestern University in Chicago, Illinois, found out. Kibbe started her career using a K08 grant from the National Heart, Lung, and Blood Institute (NHLBI) in Bethesda, Maryland. She spends two days a week working in the operating theatre; for the remaining three days she runs a lab focused on the interaction of vascular implants with the body, and the development of therapeutic materials. Her grant required that she spend at least 75% of her time on research, yet provided only $75,000 in salary support — not nearly enough to cover three-quarters of a year of surgeon's pay at a university hospital, which often far exceeds $100,000. "Who is going to come up with the rest of the salary?" she asks. Kibbe was lucky; she managed to win a matching grant from the Society for Vascular Surgery in Chicago to cover the cost inequities between the NHLBI grant cap and her salary. "The matching grants make all the difference in the world," she says. Grant salaries vary across NIH institutes. Although NHLBI caps its grantees' pay at $75,000, others allow salaries up to the current legislative cap of $199,700. A spokesperson for the NIH office of extramural research says that the physician's institution is expected to supplement the awardee's salary as an "important indication of its commitment to the development of the physician-scientist's career".

It takes real commitment and great mentoring and support to resist the pressures to spend too much time in the clinic, says Nancy Andrews, dean of the Duke University School of Medicine in Durham, North Carolina, and former head of the MD–PhD programme at Harvard University and the Massachusetts Institute of Technology in Cambridge.

Women, especially, seem vulnerable during early-career transitions, says Schafer. Although they represent a growing fraction of the MD–PhD pool, they are less likely than men to remain physician-scientists, possibly owing to a lack of role models, or to a perception that such a career is incompatible with having children. "Unless major changes occur in academe to make research careers more attractive to women physicians, the overall pool of competitive physician-scientist prospects will continue to decline in the job market," he says.

Physician-scientists have historically been in high demand, and that demand should continue, says Schafer. A 2007 AAMC survey[2] suggests that there are plenty of open positions for physician clinical investigators, with a little more than half of all departments surveyed unable to fill all their positions, and a quarter of positions overall going unfilled. According to Lawrence Brass, director of the Medical Scientist Training Program at the University of Pennsylvania, only about 550 students enter MD–PhD programmes in the United States each year. "That's not enough to sustain the number of physician-scientists working now, let alone expand it," he says. Although this is bad news for medical schools and research universities, it could be good for soon-to-be physician-scientists, who will find themselves very employable once they graduate.

Yet economic realities are adding new pressures to the physician aspect of the job. With grant dollars tightening and overall economics worsening, it's harder to justify spending 'unbillable' time on research as opposed to in the clinic. "I do have a fear that within the new era of health care we are in, that it will be more challenging for medical centres to find the margin to support people doing research and not clinical work," says Steven Gabbe, chief executive of the Ohio State University Medical Center in Columbus, and former chair of the AAMC's Clinical Research Task Force II, which authored a report on promoting translational and clinical science[3].



***"The matching grants make all the difference in the world."***
Melina Kibbe

Still, says Gabbe, there will "always be a place for the creative, well-trained physician-scientist to find a faculty position". And for those who secure such positions, the rewards can be substantial. "As a physician-scientist, you really get to craft your career, if you are thoughtful about it," says Brass. "I get to discover new things, take care of patients with problems that interest me, and work with students who are eager and bright. That sounds like a great job description, doesn't it?" ■

**Jeffrey Perkel** *is a freelance writer based in Pocatello, Idaho.*

1. Association of American Medical Colleges *2010 GQ Medical School Graduation Questionnaire, All Schools Summary Report FINAL* (AAMC, 2010). Available at go.nature.com/tyer27
2. Fang, D. *et al.* Recruitment of *New Physician Investigators in Clinical Research* (AAMC, 2007).
3. Task Force II on Clinical Research *Promoting Translational and Clinical Science* (AAMC, 2006).

UNITED STATES

## International funding

A graduate fellowship will fund 48 international doctoral students in the United States and is seeking another 50 awardees in early September. Recognizing that graduate students from abroad often have trouble securing funding, the Howard Hughes Medical Institute (HHMI) in Chevy Chase, Maryland, established the programme, an annual US$2-million commitment, this year. The fellowship provides $43,000 a year for science and engineering PhD students in their third, fourth and fifth years of graduate school. Candidates, who must demonstrate creativity and innovation, are nominated by their institution. "We're looking for superstars," says Maryrose Franko, HHMI senior programme officer for graduate science education. "We expect these people will become scientific leaders."

ITALIAN ACADEMIA

## What's in a name?

Nepotism is a widespread problem in Italian academic institutions, according to a statistical analysis of professors' last names. The study — which examined a database of more than 61,000 tenured professors from 94 institutions across 370 subdisciplines — found that the diversity of last names was lower than would be expected from unbiased hirings (S. Allesina *PLoS ONE* **6,** e21160; 2011). Study author Stefano Allesina, an Italian ecologist at the University of Chicago in Illinois, says the index probably underestimates the incidence of nepotism because it identifies the problem only in father–child and sibling relationships. Allesina says he accounted for unrelated people with the same last names.

EDUCATION

## Upward physics trend

Physics PhDs conferred by US institutions rose 38% between 2004 and 2008, the sharpest upswing in more than a decade, according to the latest survey data released by the American Institute of Physics in College Park, Maryland. Patrick Mulvey, lead research associate for the institute's statistical research centre, says the increase corresponds to a rise in enrolments in the early 2000s that was driven by an uncertain international economy and by improvements in the undergraduate culture of physics departments. Preliminary PhD totals for 2009 and 2010 show a similar trend, Mulvey says.



© 2011 Macmillan Publishers Limited. All rights reserved

available to early-career scientists as well as seasoned practitioners (see 'Mechanisms of support'). But if they fail to win such grants, physician-scientists can be required to fall back on their medical training to see more patients, thereby diminishing their time in the lab. Down-on-their-luck physician-scientists can thus become progressively less competitive in research.

### TOUGH TRANSITION

Supporting both research and clinical priorities is not easy, as Melina Kibbe, a vascular surgeon at Northwestern University in Chicago, Illinois, found out. Kibbe started her career using a K08 grant from the National Heart, Lung, and Blood Institute (NHLBI) in Bethesda, Maryland. She spends two days a week working in the operating theatre; for the remaining three days she runs a lab focused on the interaction of vascular implants with the body, and the development of therapeutic materials. Her grant required that she spend at least 75% of her time on research, yet provided only $75,000 in salary support — not nearly enough to cover three-quarters of a year of surgeon's pay at a university hospital, which often far exceeds $100,000. "Who is going to come up with the rest of the salary?" she asks. Kibbe was lucky; she managed to win a matching grant from the Society for Vascular Surgery in Chicago to cover the cost inequities between the NHLBI grant cap and her salary. "The matching grants make all the difference in the world," she says. Grant salaries vary across NIH institutes. Although NHLBI caps its grantees' pay at $75,000, others allow salaries up to the current legislative cap of $199,700. A spokesperson for the NIH office of extramural research says that the physician's institution is expected to supplement the awardee's salary as an "important indication of its commitment to the development of the physician-scientist's career".

It takes real commitment and great mentoring and support to resist the pressures to spend too much time in the clinic, says Nancy Andrews, dean of the Duke University School of Medicine in Durham, North Carolina, and former head of the MD–PhD programme at Harvard University and the Massachusetts Institute of Technology in Cambridge.

Women, especially, seem vulnerable during early-career transitions, says Schafer. Although they represent a growing fraction of the MD–PhD pool, they are less likely than men to remain physician-scientists, possibly owing to a lack of role models, or to a perception that such a career is incompatible with having children. "Unless major changes occur in academe to make research careers more attractive to women physicians, the overall pool of competitive physician-scientist prospects will continue to decline in the job market," he says.

Physician-scientists have historically been in high demand, and that demand should continue, says Schafer. A 2007 AAMC survey[2] suggests that there are plenty of open positions for physician clinical investigators, with a little more than half of all departments surveyed unable to fill all their positions, and a quarter of positions overall going unfilled. According to Lawrence Brass, director of the Medical Scientist Training Program at the University of Pennsylvania, only about 550 students enter MD–PhD programmes in the United States each year. "That's not enough to sustain the number of physician-scientists working now, let alone expand it," he says. Although this is bad news for medical schools and research universities, it could be good for soon-to-be physician-scientists, who will find themselves very employable once they graduate.

Yet economic realities are adding new pressures to the physician aspect of the job. With grant dollars tightening and overall economics worsening, it's harder to justify spending 'unbillable' time on research as opposed to in the clinic. "I do have a fear that within the new era of health care we are in, that it will be more challenging for medical centres to find the margin to support people doing research and not clinical work," says Steven Gabbe, chief executive of the Ohio State University Medical Center in Columbus, and former chair of the AAMC's Clinical Research Task Force II, which authored a report on promoting translational and clinical science[3].



***"The matching grants make all the difference in the world."***
Melina Kibbe

Still, says Gabbe, there will "always be a place for the creative, well-trained physician-scientist to find a faculty position". And for those who secure such positions, the rewards can be substantial. "As a physician-scientist, you really get to craft your career, if you are thoughtful about it," says Brass. "I get to discover new things, take care of patients with problems that interest me, and work with students who are eager and bright. That sounds like a great job description, doesn't it?" ■

**Jeffrey Perkel** *is a freelance writer based in Pocatello, Idaho.*

1. Association of American Medical Colleges *2010 GQ Medical School Graduation Questionnaire, All Schools Summary Report FINAL* (AAMC, 2010). Available at go.nature.com/tyer27
2. Fang, D. *et al.* Recruitment of *New Physician Investigators in Clinical Research* (AAMC, 2007).
3. Task Force II on Clinical Research *Promoting Translational and Clinical Science* (AAMC, 2006).

UNITED STATES

## International funding

A graduate fellowship will fund 48 international doctoral students in the United States and is seeking another 50 awardees in early September. Recognizing that graduate students from abroad often have trouble securing funding, the Howard Hughes Medical Institute (HHMI) in Chevy Chase, Maryland, established the programme, an annual US$2-million commitment, this year. The fellowship provides $43,000 a year for science and engineering PhD students in their third, fourth and fifth years of graduate school. Candidates, who must demonstrate creativity and innovation, are nominated by their institution. "We're looking for superstars," says Maryrose Franko, HHMI senior programme officer for graduate science education. "We expect these people will become scientific leaders."

ITALIAN ACADEMIA

## What's in a name?

Nepotism is a widespread problem in Italian academic institutions, according to a statistical analysis of professors' last names. The study — which examined a database of more than 61,000 tenured professors from 94 institutions across 370 subdisciplines — found that the diversity of last names was lower than would be expected from unbiased hirings (S. Allesina *PLoS ONE* **6,** e21160; 2011). Study author Stefano Allesina, an Italian ecologist at the University of Chicago in Illinois, says the index probably underestimates the incidence of nepotism because it identifies the problem only in father–child and sibling relationships. Allesina says he accounted for unrelated people with the same last names.

EDUCATION

## Upward physics trend

Physics PhDs conferred by US institutions rose 38% between 2004 and 2008, the sharpest upswing in more than a decade, according to the latest survey data released by the American Institute of Physics in College Park, Maryland. Patrick Mulvey, lead research associate for the institute's statistical research centre, says the increase corresponds to a rise in enrolments in the early 2000s that was driven by an uncertain international economy and by improvements in the undergraduate culture of physics departments. Preliminary PhD totals for 2009 and 2010 show a similar trend, Mulvey says.

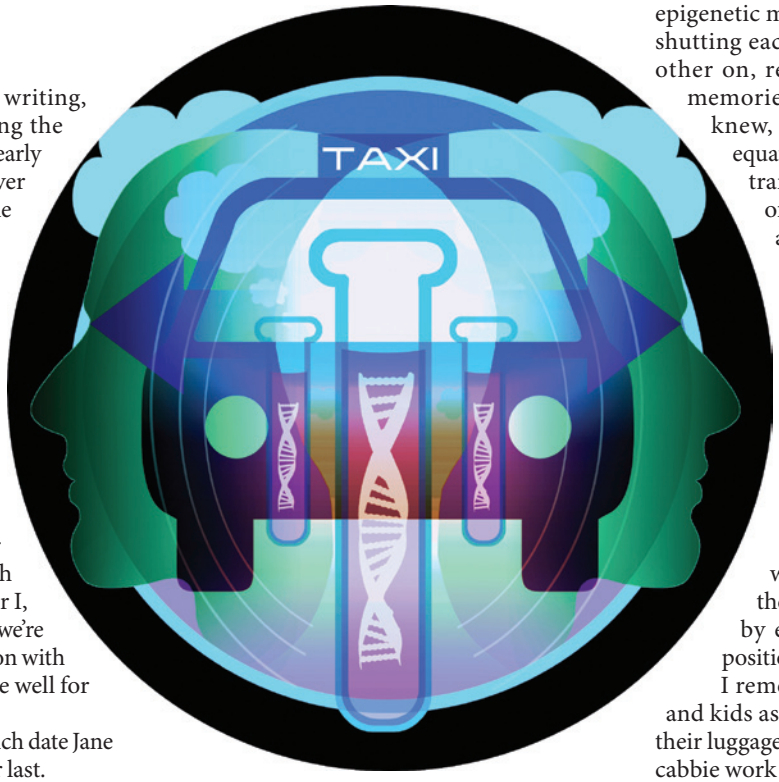


© 2011 Macmillan Publishers Limited. All rights reserved

# HER NAME WAS JANE

*Remembrance day.*

**BY JOSES HO**




JACEY

My name is Noah.

I'm going to start writing, because I'm following the only piece of advice that I can clearly recall. I heard it from a cab driver as I rode in his taxi. He said: "The people you write about should be the ones who you dream about."

Jane and I, as I remember, might have been chatting about our favourite authors, something along those lines, I'm not sure. We both looked up at the taxi driver's eyes, dimly reflected through the rear-view mirror.

"Do you write stories in your spare time?" I can't recall which one of us asked him that, Jane or I, but I remember thinking: "Now we're getting drawn into a conversation with a stranger. This might not bode well for a first date."

To be honest, I'm not sure which date Jane and I were on. I know it was our last.

"I used to write." He continued looking at us through the mirror as we sped along. "Non-fiction, though. Papers. Quite a few. Scientific ones."

His eyes turned back to the road, and he said: "It wasn't too long ago either. Two or three years. Or maybe two or three months."

"What do you mean?" He turned his face back, and I think his eyes met mine.

"Do you really want to know?"

I said, half-laughing: "Well, now you have to tell us, right?" I turned slightly to look at Jane, but I can't recall what her expression was.

The taxi driver turned back, and said: "Well, then I guess I'll show you." I don't think he said anything else for the rest of the journey. Jane and I probably just sat in awkward silence until we reached her place, and I don't think I bothered to talk to the driver as he drove me back to mine.

And, as I recall, that night the dreams started.

Those first dreams, I remember, were of my childhood — happy days with my older sister Karen. We would hide from the torrential rain, squeezing into our dollhouse, giggling like the pre-adolescent girls we were as Mother screamed over the downpour: "Karen! Jane! Come back inside!" The funny thing is, now that I think about it, I've never had a sister. I've never had a dollhouse in any garden. I don't know anyone named Karen, either. I woke up that morning terrified, in a bedroom I knew was mine but didn't recognize. I remember shattering the bathroom mirror, because the reflection wasn't Jane's, or Rob's.

Rob — he's the taxi driver, the most overqualified cabbie in the world. He never told us in the taxi, but I know his name because the next night, I dreamt I was Rob. I was in Paris, listening to the lecture that started it all. The speaker might have been a Nobel laureate, but I can't recall his name. But I can see, even now, the quantum field equations he had, leaping off the PowerPoint slide. He had a glorious, wonderful theory. He hypothesized that DNA could, if placed in a low-frequency electric field, be teleported between two test tubes. I remember most of the audience walking out halfway, shocked and stunned at the quackery he was proposing.

And I remember strolling down the Champs-Elysées, lost in my thoughts. The biological basis of memory, I reasoned, could be explained by epigenetic modifications to DNA. Genes shutting each other down, turning each other on, rewiring the brain, creating memories and erasing others. And I knew, with a few tweaks to those equations I had just seen, you could transmit the quantum signature of one neuronal epigenome to another. In other words, you could clone memories.

The thing is, if you check my passport, I've never been to France in my entire life. And I don't know what the hell the word 'epigenetic' means. And I can remember flying back and living in my lab for the next few months, despite the constant warnings from the institute that they would kick me out. And when they did, I recall being shunned by every university I sought a position with.

I remember screaming at my wife and kids as they left the apartment with their luggage, insisting that this night-shift cabbie work would tide us over, don't give up on me now, please, I'm this close to making it work. And then I see myself, waking up, hungover again, stumbling down to the taxi and rewiring the coils into the chassis, knowing there was only one way to find out if it did work.

Last night I dreamt that there's a cabal of cabbies, driving electromagnetic taxis around the city, secretly homogenizing the histories of random strangers. Because driving a taxi's the easiest way to get people to climb into a 7 Hz field for at least 15.3 minutes. That's all the time needed to swap your memories with someone else.

You see, there's only one way to know which of your memories are real. You compare them with newer dreams, ones you're sure you've never had before. So I'm taking my advice, writing all of this down, so that I can sort out all these dreams. Because I'm not sure, anymore, which of these memories are actually mine.

I don't see her in my dreams anymore — and I'm not even sure if her name was Jane. ■

**NATURE.COM**
Follow Futures on Facebook at:
go.nature.com/mtoodm

**Joses Ho** *is pursuing a DPhil at the University of Oxford, UK. He was born and bred in Singapore. Follow him on Twitter at www.twitter.com/jacuzzijo.*



© 2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10324

# α–Synuclein occurs physiologically as a helically folded tetramer that resists aggregation

Tim Bartels[1], Joanna G. Choi[1] & Dennis J. Selkoe[1]

**Parkinson's disease is the second most common neurodegenerative disorder[1,2]. Growing evidence indicates a causative role of misfolded forms of the protein α-synuclein in the pathogenesis of Parkinson's disease[3,4]. Intraneuronal aggregates of α-synuclein occur in Lewy bodies and Lewy neurites[5], the cytopathological hallmarks of Parkinson's disease and related disorders called synucleinopathies[4]. α-Synuclein has long been defined as a 'natively unfolded' monomer of about 14 kDa (ref. 6) that is believed to acquire α-helical secondary structure only upon binding to lipid vesicles[7]. This concept derives from the widespread use of recombinant bacterial expression protocols for *in vitro* studies, and of overexpression, sample heating and/or denaturing gels for cell culture and tissue studies. In contrast, we report that endogenous α-synuclein isolated and analysed under non-denaturing conditions from neuronal and non-neuronal cell lines, brain tissue and living human cells occurs in large part as a folded tetramer of about 58 kDa. Several methods, including analytical ultracentrifugation, scanning transmission electron microscopy and *in vitro* cell crosslinking confirmed the occurrence of the tetramer. Native, cell-derived α-synuclein showed α-helical structure without lipid addition and had much greater lipid-binding capacity than the recombinant α-synuclein studied heretofore. Whereas recombinantly expressed monomers readily aggregated into amyloid-like fibrils *in vitro*, native human tetramers underwent little or no amyloid-like aggregation. On the basis of these findings, we propose that destabilization of the helically folded tetramer precedes α-synuclein misfolding and aggregation in Parkinson's disease and other human synucleinopathies, and that small molecules that stabilize the physiological tetramer could reduce α-synuclein pathogenicity.**

To identify the native state of α-synuclein in cells while avoiding the potential breakdown of physiological assemblies by detergents, we initially used native gel electrophoresis. α-Synuclein is expressed endogenously in many cell types, so we chose to analyse the dopaminergic human neuroblastoma line, M17D (ref. 8) and the commonly used cell lines HEK293, HeLa and COS-7. Each of these cell lines predominantly contained a non-denatured α-synuclein-immunoreactive species migrating in blue native–polyacrylamide gel electrophoresis (BN–PAGE) at ~45–50 kDa (Fig. 1a, lanes 1–4). Because these initial results suggested an apparently stable oligomeric form under native conditions, we next probed the endogenous state of α-synuclein in normal brain. The frontal cortex of wild-type mice also revealed a ~45–50 kDa form of endogenous α-synuclein as the main species in the buffer-soluble fraction (Fig. 1a, lane 6).

To assess the state of endogenous α-synuclein in living human cells, we examined freshly collected red blood cells (RBCs), which were recently found to have high α-synuclein expression[9]. Human RBCs contained a ~45–50 kDa α-synuclein immunoreactive band on BN–PAGE (Fig. 1a, lane 5). As a second non-denaturing gel system that precludes effects of the Coomassie dye used in BN–PAGE, we performed clear native–PAGE (CN–PAGE)[10]. The main α-synuclein species in all samples migrated at ~55–60 kDa, suggesting a tetramer (theoretical mass of monomer = 14,460 Da) (Fig. 1b, lanes 1–6). The better resolution of CN–PAGE without Coomassie dye also revealed small amounts of apparent monomers running below the 14 kDa molecular weight marker (Fig. 1b, lanes 1–4, 6) and distinguished the small differences in amino acid length of the human and mouse α-synuclein monomers and putative tetramers (Fig. 1b, lane 6). The endogenous ~55–60 kDa species was detected by monoclonal α-synuclein antibodies Syn1, Syn211 and LB509 and polyclonal antibody C20 in both native gel systems.

Because the migration of a protein on BN– or CN–PAGE does not depend solely on its mass but also on its conformation and charge, we used *in vitro* crosslinking to preserve the assembled state of the putative α-synuclein oligomer, followed by denaturing SDS–PAGE. We observed SDS-stable α-synuclein bands migrating at the expected positions of a tetramer (~55 kDa) and non-crosslinked monomer in all

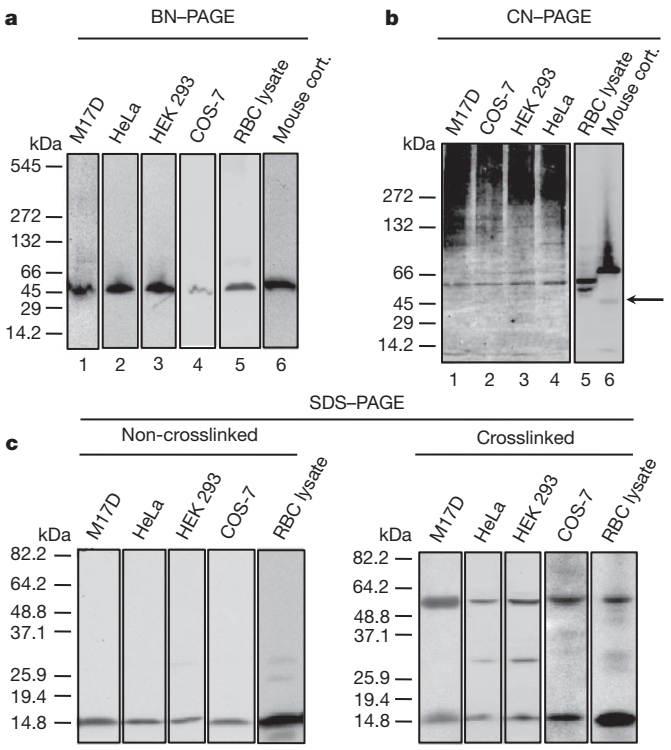



**Figure 1 | Western blot analysis of lysates of M17D, HeLa, HEK293 and COS-7 cells, mouse cortex and human RBCs probed for endogenous α-synuclein.** **a**, BN–PAGE. **b**, CN–PAGE. The band just below the main ~55–60 kDa RBC species (lane 6) may represent an alternatively spliced form of α-synuclein. Arrow marks a possible dimeric species. **c**, Left, SDS–PAGE/western blot (antibody C20) analysis of cell lysates without crosslinking. Right, proteins were crosslinked in intact living cells with membrane permeable disuccinimidyl suberate (DSS) (M17D, HeLa, HEK 293, COS-7) or in RBC lysate with water soluble bis(sulphosuccinimidyl) suberate ($BS^3$) and then run on SDS–PAGE.

[1]Center for Neurologic Diseases, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA.



©2011 Macmillan Publishers Limited. All rights reserved

cells, plus some putative dimer in the HeLa, HEK and red blood cells (right panel of Fig. 1c). This *in vivo* crosslinking supports the existence of native tetramers in cells. We performed two-dimensional gel analysis after the *in vivo* crosslinking; that is, isoelectric focusing (IEF) to separate proteins by charge in a pH gradient followed by denaturing SDS–PAGE. The higher-migrating α-synuclein species in the crosslinked RBC lysates had the same p$K_a$ as monomers, within the limits of IEF resolution (Supplementary Fig. 1), consistent with their being homo-oligomers.

Next, we developed a non-denaturing method to purify native α-synuclein from soluble RBC lysates (see Methods and *Nature* Protocol Exchange (http://www.nature.com/protocolexchange/protocols/2136)). This allowed us to estimate the mass of native α-synuclein based on distinct measurement principles that are not affected by protein conformation, unlike gel electrophoresis. Scanning transmission electron microscopy (STEM) is useful for measuring the masses of purified, non-covalently bonded complexes that may not resist ionization during mass spectrometry[11,12]. STEM images of α-synuclein purified under non-denaturing conditions from human RBCs (Supplementary Fig. 2) yielded a homogenous distribution of roughly spherical particles measuring ~3.0–3.5 nm in diameter (Fig. 2a). Unbiased automatic sampling of 1,000 particles gave a size distribution pattern with a peak at ~55 kDa (Fig. 2b). Importantly, we next applied sedimentation equilibrium analytical ultracentrifugation (SE-AUC), a technique commonly used to establish the oligomeric state of native proteins independent of their conformation. SE-AUC analysis of purified, native RBC α-synuclein performed at three different concentrations and at different rotor speeds yielded an average molecular weight of 57.8 kDa (4.78 Svedbergs), strongly supporting a tetrameric assembly state (Fig. 2c).

Numerous studies have reported conformational changes in α-synuclein, with a focus on the natively unfolded recombinant monomer undergoing a random coil to α-helix transition upon *in vitro* interaction with small lipid vesicles[7]. This change is believed to be relevant to the poorly defined physiological function of α-synuclein in cells and could potentially decrease the likelihood of its aggregation into β-sheet-rich neurotoxic assemblies[13]. Unexpectedly, we found that circular dichroism spectra of the human RBC tetramer purified under non-denaturing conditions showed two minima of mean residue ellipticity at 222 and 208 nm (Fig. 3a), characteristic of an α-helically folded protein[14]. This result is inconsistent with the common assumption that α-synuclein is natively unfolded. The addition of negatively charged, small unilamellar lipid vesicles (SUVs) did not induce a significant conformational change in the native tetramer by circular dichroism (Fig. 3a), but a random coil to α-helical conversion did occur (as reported) with recombinant monomer that had been expressed in bacteria (Fig. 3b). Incubation of the purified RBC α-synuclein tetramer with Lipidex 1000, a reagent used to strip proteins of bound lipids and fatty acids[15], did not change the conformation of the α-helical α-synuclein tetramer (Supplementary Fig. 3), suggesting that significant lipid association is not required to maintain the folded structure of cellular α-synuclein. To support this possibility, we conducted a quantitative elemental phosphate analysis[16] on the purified native α-synuclein to search for phospholipid. We obtained an average value of 0.25 mol phosphate per mol α-synuclein, making a significant presence of phospholipids on the α-helical α-synuclein purified from normal cells unlikely. Because post-translational modifications also could have an impact on the conformational differences between the native human RBC tetramer and the bacterially expressed, recombinant human monomer, we performed mass spectrometry. The recombinant protein showed a mass peak at 14,462 kDa, very close to the theoretical predicted mass of 14,460 kDa, whereas the purified erythrocyte α-synuclein showed a peak at 14,505 kDa, indicative of only an N-α-acetylation commonly present on human proteins (theoretical predicted mass = 14,502 kDa) (Supplementary Fig. 4).

To validate the above results obtained on RBC α-synuclein using a different human cell type and a different non-denaturing purification method, we isolated α-synuclein from a M17D human neuroblastoma cell line stably overexpressing wild-type human α-synuclein (3D5 cells[17]). α-Synuclein from untransfected M17D cell lysates migrated above bacterially expressed α-synuclein of confirmed random coil structure on CN–PAGE (Supplementary Fig. 5A). This higher electrophoretic migration was also true of native (α-helical) but not

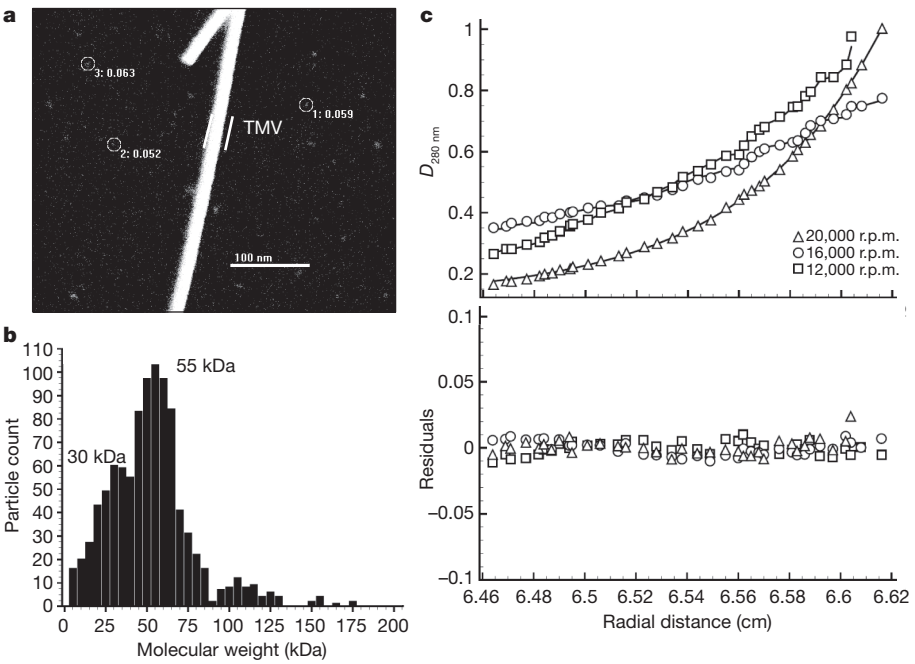



**Figure 2 | Sizing analyses of α-synuclein from human RBCs.**
**a**, Representative large-angle dark-field STEM image of purified α-synuclein from human RBCs. A few representative particles are circled. As an internal size standard, tobacco mosaic virus (TMV) helical rod was included during electron microscopy specimen preparation. **b**, Mass histogram (bin size = 5 kDa) of 1,000 automatically selected α-synuclein particles. **c**, SE-AUC of purified, native RBC α-synuclein. Top panel shows the individual experimental analyses fitting an ideal single-species model to the equilibrium data obtained at 11,612, 20,644 and 32,256*g* for 1.1 mg ml$^{-1}$ α-synuclein solution. The fitting yielded a molecular weight of 57,753 Da (standard deviation (s.d.) ± 655.199) with a root mean squared deviation of 0.004533. *D*, attenuance. Bottom panel shows an overlay of the residuals of data and theoretical fit for the three different speeds.



©2011 Macmillan Publishers Limited. All rights reserved

denatured (random coil) purified RBC α-synuclein (Supplementary Fig. 5B). After α-synuclein was purified from the stably transfected 3D5 cell line or from RBCs, the two differentially purified and α-helically folded (by circular dichroism) cellular proteins co-migrated at ~55–60 kDa on CN–PAGE, as expected (Supplementary Fig. 6). Unbiased, automated STEM measurements of 3,000 particles revealed that the 3D5 neuroblastoma cells contained α-synuclein tetramers of closely similar estimated molecular weight (peak mass ~55 kDa) to those of the RBC α-synuclein (Supplementary Fig. 7; compare to Fig. 2b). Circular dichroism spectroscopy revealed the purified 3D5 cell α-synuclein to have two minima of mean residue ellipticity at 222 and 208 nm (Supplementary Fig. 8). To further exclude artefacts arising during purification of cellular α-synuclein such as adventitious association of biomolecules (for example, cellular lipids not removed by Lipidex 1000) that artificially fold the protein, we repeated our experiments with the 3D5 parental line M17D, which has only low levels of endogenous α-synuclein. We added ('spiked') bacterially expressed recombinant human monomer onto the M17D cells before performing cell lysis and the full purification, and then assayed its structural properties. This exposure to cell lysates and the purification procedure led to no induction of helical folding in the recombinant human α-synuclein (Supplementary Fig. 9), whereas simultaneously purified 3D5 cell human α-synuclein did show this conformation, supporting our conclusion that α-helically folded α-synuclein does not arise owing to artificial manipulation of the protein.

Membrane association has been viewed as a principal functional property of α-synuclein *in vitro*[7] and in living cells[18]. We searched for differential binding of recombinant monomeric human α-synuclein versus RBC tetrameric human α-synuclein to a lipid membrane using surface plasmon resonance (SPR). Because recombinant α-synuclein is reported to have preferential affinity for negatively charged lipids, especially phosphatidyl serine[7], we chose a mixed phosphatidyl choline and phosphatidyl serine (PC/PS) membrane as a model membrane. Exposure of a PC/PS membrane to cell-derived, purified native α-synuclein in a Biacore instrument produced a markedly increased resonance angle shift compared to conventional recombinant monomers at identical concentrations in solution (Fig. 3c), indicating dramatically increased lipid binding. Fitting a dilution series of α-synuclein tetramer injections to a two-state binding model (Supplementary Fig. 10) gave an apparent dissociation constant of $K_{app} = 56 \pm 61$ nM, which is about two orders of magnitude lower than values obtained for recombinant monomer in analogous SPR studies[19]. We next tested the amyloid aggregation propensity of the distinct species in a Thioflavin T (ThT) fluorescence assay. Monomeric and tetrameric α-synuclein showed very different characteristics, with samples of purified cellular α-synuclein incubated under identical conditions showing no evidence of fibril formation in a time (10 days) more than sufficient to form mature, ThT-bound fibrils from equivalent amounts of unfolded recombinant α-synuclein (Fig. 3d). Analysis of protein concentration in the solution after the 10-day incubation showed that the RBC α-synuclein was still present and soluble, making non-amyloid (that is, ThT-negative) aggregation of the tetramers unlikely. Interestingly, melting curves of purified tetrameric α-synuclein showed that heat denaturation (at 95 °C) seemed irreversible under our conditions (Supplementary Fig. 11).

Our experiments provide several independent lines of evidence that endogenous cellular α-synuclein exists in large part as an α-helically folded, ~58 kDa tetramer under native conditions. This finding is in contrast to many biophysical and biochemical studies describing α-synuclein as a natively unfolded ~14 kDa monomer. In an early study of bacterially expressed recombinant protein purified under non-denaturing conditions or with heat treatment, no conformational differences were observed, and it was concluded that α-synuclein is a natively unfolded monomer[6]. This suggests problems in generating properly folded protein in Escherichia coli, although a modified bacterial expression protocol avoiding heating and denaturants has recently been found to yield a helical α-synuclein tetramer closely resembling the species found by us in native human samples (W. Wang *et al.*, personal communication). The reasons for the conformational differences observed in these two bacterial studies are unknown. Using gel filtration on unfolded recombinant α-synuclein also showed an apparent molecular weight of ~60 kDa in some earlier studies; the data were interpreted as a decrease in mobility of the extended state of an unfolded protein in the tested matrices[6]. This suggests the possibility

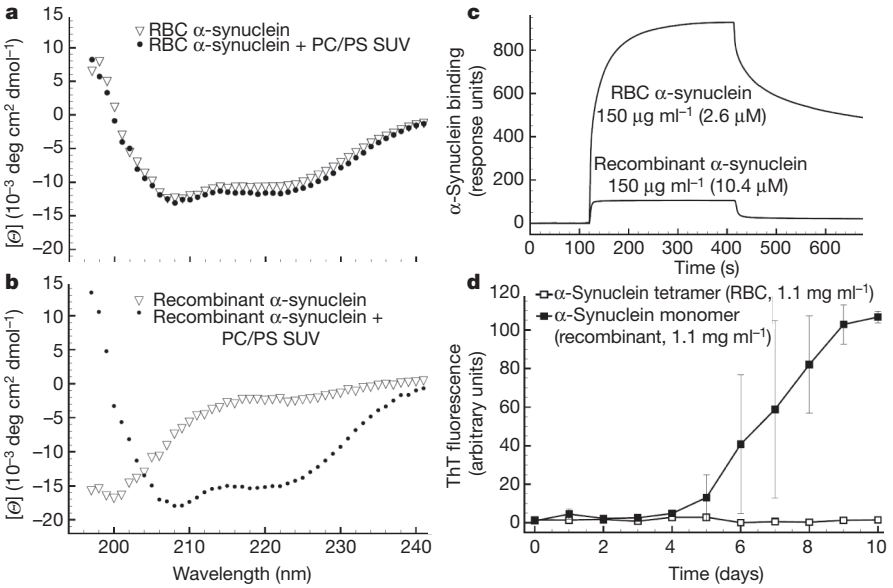



**Figure 3 | Comparative analyses of native (cell-derived) and bacterial α-synuclein.** **a**, Circular dichroism spectra of native tetrameric α-synuclein (isolated under non-denaturing conditions from human RBCs) before versus after addition of POPC/POPS SUVs (PC/PS 4:1; protein:lipid 1:500). **b**, Circular dichroism spectra of recombinant α-synuclein monomer purified from *E. coli*, alone and with addition of PC/PS SUVs (protein/lipid 1:500). **c**, SPR sensorgram of equal protein concentrations of α-synuclein recombinant monomer versus endogenous tetramer injected on a L1 chip covered with a PC/PS membrane. **d**, Amyloid-type aggregation kinetics of recombinant α-synuclein monomer versus native RBC tetramer monitored by ThT fluorescence; average values from 3 independent experiments (error bars show s.d.; some s.d. values for RBC-derived α-synuclein are smaller than the symbol size).



©2011 Macmillan Publishers Limited. All rights reserved

of a similar hydrodynamic radius for the unfolded monomer and the more compact, helically folded tetramer, making gel filtration an unreliable indicator (and therefore not used here). Our evidence for a tetrameric molecular mass of endogenous α-synuclein was particularly supported by the analytical ultracentrifugation and the unbiased STEM analysis, both of which sizing methods are not based on conformation. The STEM sizing was performed on intrinsic α-synuclein isolated from two cell types and using two distinct non-denaturing procedures.

Our apparent disagreement with most published findings on the monomeric state of α-synuclein in cells and brain tissue, usually as judged by SDS–PAGE and western blotting, can be explained by the common use of denaturing detergents. Our tetramer aggregation data (Fig. 3d) are consistent with a recent report describing non-neurotoxic, aggregation-resistant α-synuclein oligomers *in vivo*[20]. Moreover, an oligomeric species of α-synuclein (size undefined) was observed by *in vivo* fluorescence lifetime imaging in an intact cell culture model[21]. Given the close match between our observed molecular weights using SE-AUC (Fig. 2c) and STEM (Fig. 2b and Supplementary Fig. 7) and the theoretical weight of a tetramer, the detection of a tetrameric band on denaturing gels after *in vivo* crosslinking (Fig. 1c), and the IEF evidence post-crosslinking that the endogenous tetramer and dimer bands have p$K_a$ values similar to that of a monomer (Supplementary Fig. 1), we conclude that the predominant physiological species of α-synuclein in cells and brain is a helically folded tetramer, although minor and variable amounts of monomers, dimers and trimers were detected in some cell types. The closely similar properties of α-synuclein observed until now in neural cells and fresh human RBCs recommends the latter as an abundant, available source for future studies of physiological α-synuclein.

The higher lipid-binding capacity of native α-synuclein leads us to speculate that the monomer represents a not fully functional and less abundant species in normal cells. Given the much lower propensity of the native tetramer to aggregate into fibrils (Fig. 3d), it is likely that tetramers undergo destabilization before α-synuclein aggregates into abnormal oligomeric and fibrillar assemblies that can confer cytotoxicity in Parkinson's disease and other α-synucleinopathies. Hypothetically, such a mechanism could be analogous in part to transthyretin amyloidosis, in which a native metastable tetramer circulates in human plasma but can become destabilized (for example, by pathogenic missense mutations) to allow monomers to aggregate aberrantly in tissue[22]. Our identification of helically folded α-synuclein tetramers encourages the design of compounds that, like those for transthyretin[23], could kinetically stabilize native tetramers and prevent pathogenic α-synuclein aggregation as a novel treatment approach for Parkinson's disease, dementia with Lewy bodies and other synucleinopathies[24].

## METHODS SUMMARY

Native gel electrophoresis was conducted as described[10]. For crosslinking, 1–5 mM disuccinimidyl suberate was added to living cells. RBC lysates were treated analogously but using 1 mM bis(sulphosuccinimidyl) suberate. To purify α-synuclein from fresh or packed frozen RBCs, an initial 25% $(NH_4)_2SO_4$ cut followed by a 50% $(NH_4)_2SO_4$ precipitation substantially enriched α-synuclein. The resolubilized 50% pellet was injected onto a hydrophobic interaction column (HiTrap Phenyl HP; GE Healthcare) and eluted in 1 M to 0 M $(NH_4)_2SO_4$, pH 7. Alternatively, α-synuclein-overexpressing 3D5 neuroblastoma cell lysate after $(NH_4)_2SO_4$ was injected onto a 5-ml HiTrap Q HP column. A 25–500 mM NaCl (pH 8.0) gradient eluted α-synuclein at ~300 mM NaCl. α-Synuclein from both cell sources underwent a final purification step on a Superdex 75 SEC column. STEM analysis was conducted at the Brookhaven National Laboratory STEM user facility. Sedimentation equilibrium data were acquired on a Beckman XL-I analytical ultracentrifuge at speeds of 11,612, 20,644 and 32,256*g* (AN-60 Ti rotor) and protein concentrations of 0.6, 1.1 and 1.6 mg ml$^{-1}$. Circular dichroism spectroscopy for lipid-induced α-synuclein folding was conducted in the presence of 4 mM PC/PS (4:1) SUVs. SPR spectroscopy was conducted as described[19]. To quantify amyloid fibril growth, aliquots (10 μl) of purified α-synuclein were added to a 10 μM ThT solution in 10 mM glycine buffer, pH 9. ThT fluorescence was measured by exciting at 444 nm and scanning the emission wavelengths from 460–550 nm.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 15 November 2010; accepted 20 June 2011.**
**Published online 14 August 2011.**

1. Obeso, J. A. *et al.* Missing pieces in the Parkinson's disease puzzle. *Nature Med.* **16,** 653–661 (2010).
2. Gupta, A., Dawson, V. L. & Dawson, T. M. What causes cell death in Parkinson's disease? *Ann. Neurol.* **64,** S3–S15 (2008).
3. Winklhofer, K. F., Tatzelt, J. & Haass, C. The two faces of protein misfolding: gain- and loss-of-function in neurodegenerative diseases. *EMBO J.* **27,** 336–349 (2008).
4. Tong, J. *et al.* Brain α-synuclein accumulation in multiple system atrophy, Parkinson's disease and progressive supranuclear palsy: a comparative investigation. *Brain* **133,** 172–188 (2010).
5. Spillantini, M. G. *et al.* α-Synuclein in Lewy bodies. *Nature* **388,** 839–840 (1997).
6. Weinreb, P. H., Zhen, W., Poon, A. W., Conway, K. A. & Lansbury, P. T. J. NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry* **35,** 13709–13715 (1996).
7. Davidson, W. S., Jonas, A., Clayton, D. F. & George, J. M. Stabilization of α-synuclein secondary structure upon binding to synthetic membranes. *J. Biol. Chem.* **273,** 9443–9449 (1998).
8. DeTure, M. *et al.* Missense tau mutations identified in FTDP-17 have a small effect on tau-microtubule interactions. *Brain Res.* **853,** 5–14 (2000).
9. Scherzer, C. R. *et al.* GATA transcription factors directly regulate the Parkinson's disease-linked gene α-synuclein. *Proc. Natl Acad. Sci. USA* **105,** 10907–10912 (2008).
10. Wittig, I. & Schagger, H. Advantages and limitations of clear-native PAGE. *Proteomics* **5,** 4338–4346 (2005).
11. Osenkowski, P. *et al.* Cryoelectron microscopy structure of purified •-secretase at 12 Å resolution. *J. Mol. Biol.* **385,** 642–652 (2009).
12. Wall, J. S., Simon, M. N., Lin, B. Y. & Vinogradov, S. N. Mass mapping of large globin complexes by scanning transmission electron microscopy. *Methods Enzymol.* **436,** 487–501 (2008).
13. Beyer, K. Mechanistic aspects of Parkinson's disease: α-synuclein and the biomembrane. *Cell Biochem. Biophys.* **47,** 285–299 (2007).
14. Chen, Y., Yang, J. T. & Martinez, H. M. Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion. *Biochemistry* **11,** 4120–4131 (1972).
15. Sharon, R. *et al.* α-Synuclein occurs in lipid-rich high molecular weight complexes, binds fatty acids, and shows homology to the fatty acid-binding proteins. *Proc. Natl Acad. Sci. USA* **98,** 9110–9115 (2001).
16. Chen, P. S., Toribara, T. Y. & Warner, H. Microdetermination of phosphorus. *Anal. Chem.* **28,** 1756–1758 (1956).
17. Ko, L. W., Ko, H. H., Lin, W. L., Kulathingal, J. G. & Yen, S. H. Aggregates assembled from overexpression of wild-type α-synuclein are not toxic to human neuronal cells. *J. Neuropathol. Exp. Neurol.* **67,** 1084–1096 (2008).
18. McLean, P. J., Kawamata, H., Ribich, S. & Hyman, B. T. Membrane association and protein conformation of α-synuclein in intact neurons. Effect of Parkinson's disease-linked mutations. *J. Biol. Chem.* **275,** 8812–8816 (2000).
19. Smith, D. P. *et al.* Formation of a high affinity lipid-binding intermediate during the early aggregation phase of α-synuclein. *Biochemistry* **47,** 1425–1434 (2008).
20. Tsika, E. *et al.* Distinct region-specific α-synuclein oligomers in A53T transgenic mice: implications for neurodegeneration. *J. Neurosci.* **30,** 3409–3418 (2010).
21. Klucken, J., Outeiro, T. F., Nguyen, P., McLean, P. J. & Hyman, B. T. Detection of novel intracellular α-synuclein oligomeric species by fluorescence lifetime imaging. *FASEB J.* **20,** 2050–2057 (2006).
22. Quintas, A., Saraiva, M. J. M. & Brito, R. M. M. The tetrameric protein transthyretin dissociates to a non-native monomer in solution. *J. Biol. Chem.* **274,** 32943–32949 (1999).
23. Connelly, S., Choi, S., Johnson, S. M., Kelly, J. W. & Wilson, I. A. Structure-based design of kinetic stabilizers that ameliorate the transthyretin amyloidoses. *Curr. Opin. Struct. Biol.* **20,** 54–62 (2010).
24. Lansbury, P. T. & Lashuel, H. A. A century-old debate on protein aggregation and neurodegeneration enters the clinic. *Nature* **443,** 774–779 (2006).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** Mass measurements were carried out at the Brookhaven National Laboratory STEM facility, a user facility supported by the US Department of Energy. We are grateful to D. Walker and J. Anderson (Elan Pharmaceuticals) for conducting mass spectrometry of our purified α-synuclein samples and for comments. We thank X. Simon and I. Perovic (Brandeis University) for their assistance with the AUC and phosphate analyses. Supported by NIH grants NS051318 and NS038375 (D.J.S.). We thank our colleagues at the Center for Neurologic Diseases for many discussions.

**Author Contributions** All experiments were planned by T.B. and D.J.S. and conducted by T.B. and J.G.C. The manuscript was prepared by T.B. and D.J.S.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.J.S. (dselkoe@rics.bwh.harvard.edu).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Materials.** Recombinant human α-synuclein was bought from Anaspec. Recombinant human transthyretin was provided by I. Rappley and J. Kelly (Scripps Research Institute). HEK, COS-7 and HeLa cells were cultured in DMEM with 10% fetal bovine serum, penicillin (100 U $ml^{-1}$), streptomycin (100 μg $ml^{-1}$) and L-glutamine (2 mM). For M17D and 3D5 human neuroblastoma cells, standard DMEM was supplemented with 400 μg $ml^{-1}$ G418 and 1 μg $ml^{-1}$ puromycin. Frontal cortex was obtained from wild-type mice aged 4–9 months. α-Synuclein immunoblotting used antibodies C20 (1:1,000, Santa Cruz), LB509 (1:400, Santa Cruz), Syn211 (1:200, Santa Cruz) and Syn1 (1:2,000, BD).

**Lipid preparation.** PC/PS SUVs (30 nm) of 80% 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and 20% 1-palmitoyl-2-oleoyl-sn-glycero-3-[phospho-L-serine] (POPS) (Avanti Polar Lipids) were prepared in 10 mM sodium phosphate, pH 7.4, by sonication.

**Crosslinking.** Cells were detached and incubated at room temperature (22 °C) for 30 min in DSS crosslinker (1–5 mM), then quenched with 1 M Tris buffer, pH 7.4, for 15 min at room temperature. Human RBC lysates were treated analogously with 1 mM $BS^3$ (Pierce) to covalently crosslink lysine residues.

**BN–PAGE and CN–PAGE.** For BN–PAGE, samples were run on 4–16% Bis-Tris BN–PAGE gels (Invitrogen) at 200 V and room temperature. The cathode buffer was 50 mM tricine, 15 mM Bis-Tris, 0.02% Brilliant Blue G (Serva), pH 7.0; the anode buffer was 50 mM Bis-Tris pH 7.0. CN–PAGE was conducted identically to BN–PAGE, but Coomassie blue was omitted from the sample and the cathode buffer. Electroblotting of protein on PVDF membranes (0.45 μm pore size) was conducted at 400 mA for 2 h. For molecular weight estimation three different molecular weight marker were loaded on each gel (Sigma Non-denaturing, 108K6408, Invitrogen Native Mark, LC0725 GE Healthcare HMW Native Marker Kit, 17-0445-01).

**IEF two-dimensional PAGE.** We used the IPGphor IEF system (GE Healthcare). Lysates were heated at 65 °C overnight and brought to 200 μl with sample rehydration buffer (7 M urea, 2 M thiourea, 2% Chapso, 0.5% IPG buffer (GE Healthcare, bromophenol blue) and applied on an 11 cm 1D Ready-strip (Bio-Rad) with a pH gradient of 4–7. Sample was rehydrated for 16 h followed by IEF at 500 V for 30 min, then 1,000 V for another 30 min, and then 8,000 V for 3.5 h. The 1D strip was then applied to a precast NuPAGE ZOOM 4–12% Bis-Tris gel (Invitrogen) and run at 200 V.

**Purification of α-synuclein from human RBCs.** Freshly collected and washed RBCs were resuspended in a threefold volume of ACK lysing buffer (Lonza). $(NH_4)_2SO_4$ to a final concentration of 25% was added and incubated at 4 °C for 30 min. The lysate was centrifuged (20,000g, 20 min), and the supernatant brought up to 50% $(NH_4)_2SO_4$. The pellet was washed several time in 55% $(NH_4)_2SO_4$ to remove excess haemoglobin. The sample was centrifuged at 20,000g for 20 min and the pellet resolubilized in a 50-fold volume of 50 mM phosphate buffer, pH 7.0, 1 M $(NH_4)_2SO_4$. Five millilitres of the resultant solution were injected onto a 5-ml HiTrap phenyl hydrophobic interaction column (GE Healthcare) equilibrated with 50 mM phosphate buffer, pH 7.0, 1 M $(NH_4)_2SO_4$. α-Synuclein was eluted with a 1 M to 0 M $(NH_4)_2SO_4$ gradient in 50 mM phosphate buffer, pH 7.0 (α-synuclein eluted at ~0.75 M $(NH_4)_2SO_4$). For anion exchange purification of RBC α-synuclein, we used the protocol used for neuroblastoma cells (see later), but the first run of RBC lysate sometimes showed low binding of α-synuclein and contamination by plasma transthyretin. In these cases we discarded the first eluate and used the flow-through for a second run, which showed significantly higher binding and subsequent purity. As a third alternative to HIC and AX, an XK 16/100 column packed with activated thiolpropyl Sepharose 6B gel media was used (Supplementary Fig. 5B) (binding buffer, PBS; flow rate, 0.2 ml $min^{-1}$). In this case, the flow-trough contained α-synuclein and was processed further. The column was regenerated by eluting bound protein with 5 column volumes of binding buffer with 25 mM dithiothreitol, and reactivated with 1.5 mM dipyridyl sulphide in 50 mM borate buffer pH 8.0. The final solution was concentrated (Amicon Ultra centrifugal filter units, MWCO 10,000, Millipore) and further purified via gel filtration.

**Purification of α-synuclein from human neuroblastoma cells.** 3D5 cells (α-synuclein stables) and their parental M17D cells were scraped from the plates, washed in PBS and lysed by sonication. A $(NH_4)_2SO_4$ precipitation was conducted as described earlier, the 50% $(NH_4)_2SO_4$ pellet was taken up in 20 mM tris buffer, pH 8.0, 25 mM NaCl. The sample was injected onto a 5-ml HiTrap Q HP anion exchange column, equilibrated with 20 mM Tris buffer, pH 8.0, 25 mM NaCl. α-Synuclein was eluted from the column with a 25–500 mM NaCl gradient in 20 mM Tris buffer, pH 8.0. α-Synuclein was eluted at ~300 mM NaCl. The column was regenerated with 1 M NaCl in 20 mM tris buffer, pH 8.0. The final solution was concentrated (Amicon Ultra centrifugal filter units, MWCO 10,000, Millipore) and further purified via gel filtration. For the addition ('spiking') of exogenous recombinant α-synuclein monomers, bacterially expressed α-synuclein was added to the scraped M17D cell pellet and the purification scheme conducted as just described.

**Gel filtration.** Aliquots (250 μl) were injected onto either a Superdex 75 (10/300 GL), Superdex 200 (10/300 GL) or a Superose 12 (10/300 GL) column (GE Healthcare) at 4 °C and eluted with 50 mM ammonium acetate, pH 8.5. For size estimation, a gel filtration standard (Bio-Rad, cat. no. 151-1901) was run on each column, and the calibration curve was obtained by semi-logarithmic plotting of molecular weight versus elution volume divided by void volume.

**STEM.** STEM was carried out at the Brookhaven National Laboratory STEM user facility with 100 μl of sample at a concentration of 300 μg $ml^{-1}$ in 50 mM ammonium acetate, pH 7.4, and diluted to find the appropriate concentration for a homogenous particle distribution. Tobacco mosaic virus (TMV) rods were included during specimen preparation as an internal sizing standard.

**Circular dichroism spectroscopy.** Circular dichroism spectra were obtained using an Aviv Biomedical spectrometer (model 410) in the presence or absence of 4 mM PC/PS SUVs. The spectral contributions of buffer and SUVs were subtracted. Data are reported as mean residue ellipticities measured at 20 °C and a pathlength of 0.1 mm.

**Lipidex 1000 treatment.** 10% (w/v) Lipidex 1000 (Perkin Elmer) was washed with 50% methanol-ultra pure water and added to a 100 μM solution of purified α-synuclein from RBCs. The samples were stirred overnight at 37 °C, and α-synuclein was purified from that mixture via size exclusion chromatography.

**SPR.** All lipid binding experiments were performed at 20 °C on a BIACORE 3000 apparatus using the L1 sensor chip (Biacore AB). The running buffer was 10 mM sodium phosphate, pH 7.4. SUVs were applied to the sensor chip surface at a flow rate of 10 μl $min^{-1}$ in the presence of 0.1 mM NaCl. Injections were done at a flow rate of 10 μl $min^{-1}$ with 50 μl sample volume. Apparent $K_d$ values were calculated from equilibrium data of several dilution series, collected at 300–320 s.

**ThT binding.** To detect amyloid fibril growth, a discontinuous assay was used. Aliquots (10 μl) were removed from each purified α-synuclein sample (lyophilized from 50 mM ammonium acetate, pH 7.4, and agitated at 37 °C at a concentration of 75 μM in 20 mM Bis-Tris propane, 100 mM LiCl, pH 7.4) and added to 2 ml of a 10 μM ThT solution in 10 mM glycine buffer, pH 9. Fluorescence was directly quantified on a Varian Eclipse fluorescence spectrophotometer at 20 °C by exciting at 444 nm and scanning the emission wavelengths from 460–550 nm with slit widths set at 5 nm (PMT at 750 V).

**Quantitative phosphate analysis.** Samples (2 × 15 μl and 2 × 30 μl of 1 mg $ml^{-1}$ α-synuclein in 50 mM ammonium acetate, pH 7.4) were placed at the bottom of glass test tubes, 225 μl of 8.9 N $H_2SO_4$ (in deionized water) was added, and the mixture was heated for 25 min at 200–215 °C. Next, 75 μl $H_2O_2$ was added to all tubes at room temperature. After heating for 30 min at 200–215 °C, 1.95 ml deionized water and then 0.25 ml 2.5% ammonium molybdate(VI) tetrahydrate solution (in deionized water) were added at room temperature. After addition of 0.25 ml 10% ascorbic acid solution (in deionized water), the tubes were heated for 7 min at 100 °C, and samples were allowed to cool to room temperature. Absorbance at 820 nm was measured, and phosphate concentration calculated using a calibration curve obtained from 7 phosphate standard solutions ranging from 0–50 nmol phosphate (Sigma-Aldrich).

**SE-AUC.** AUC experiments were performed in a Beckman Optima XL-I analytical ultracentrifuge. Sedimentation equilibrium experiments were carried out at purified α-synuclein protein concentrations of 1.6, 1.1 and 0.6 mg $ml^{-1}$ in 50 mM ammonium acetate, pH 8.5. The experiments were performed at 20 °C at 11,612, 20,644 and 32,256*g* (AN-60 Ti rotor), and data were collected at 278 nm. The software SEDPHAT (version 6.5) was used to calculate the *M* and *s* of the species present in equilibrium in the samples. For molecular weight analysis, we used the model 'Species Analysis' available in the SEDPHAT program with RI noise baseline correction. Analysis was performed for each protein concentration separately, and the molecular weight determined from the average obtained for the analyses of the 3 protein concentrations. The average errors and standard deviations were calculated using Monte-Carlo simulation, with 1,000 iterations and a confidence level of 0.68.

©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10328

# Chronological evidence that the Moon is either young or did not have a global magma ocean

Lars E. Borg[1], James N. Connelly[2], Maud Boyet[3] & Richard W. Carlson[4]

**Chemical evolution of planetary bodies, ranging from asteroids to the large rocky planets, is thought to begin with differentiation through solidification of magma oceans many hundreds of kilometres in depth[1–3]. The Earth's Moon is the archetypical example of this type of differentiation. Evidence for a lunar magma ocean is derived largely from the widespread distribution, compositional and mineralogical characteristics, and ancient ages inferred for the ferroan anorthosite (FAN) suite of lunar crustal rocks. The FANs are considered to be primary lunar flotation-cumulate crust that crystallized in the latter stages of magma ocean solidification. According to this theory, FANs represent the oldest lunar crustal rock type[2–4]. Attempts to date this rock suite have yielded ambiguous results, however, because individual isochron measurements are typically incompatible with the geochemical make-up of the samples, and have not been confirmed by additional isotopic systems[5–9]. By making improvements to the standard isotopic techniques, we report here the age of crystallization of FAN 60025 using the $^{207}$Pb–$^{206}$Pb, $^{147}$Sm–$^{143}$Nd and $^{146}$Sm–$^{142}$Nd isotopic systems to be 4,360 ± 3 million years. This extraordinarily young age requires that either the Moon solidified significantly later than most previous estimates or the long-held assumption that FANs are flotation cumulates of a primordial magma ocean is incorrect. If the latter is correct, then much of the lunar crust may have been produced by non-magma-ocean processes, such as serial magmatism[10].**

To constrain the age of a FAN using multiple chronometers we (1) obtained a unique mafic-mineral-rich sample of FAN 60025 found by examining the lunar collection at the Johnson Space Center, (2) developed a new methodology for Pb–Pb chronometry in which samples are washed numerous times to remove Pb contamination, and (3) produced a 99.988% pure $^{150}$Nd spike that allowed parent/daughter ratios and isotopic compositions for both $^{146}$Sm–$^{142}$Nd and $^{147}$Sm–$^{143}$Nd chronometry to be measured in the same mineral fractions. This FAN was selected because it is coarse-grained and contains very low abundances of siderophile elements, indicating that it has been minimally altered by impact processes[11,12]. Furthermore, it has experienced a minimal flux of thermal neutrons, eliminating the need for corrections on Sm and Nd isotopic measurements[5,13]. Isotopic measurements were completed using large (about 500 mg) fractions of mafic minerals (mostly pyroxene) and plagioclase (see Supplementary Information).

The Pb–Pb and $^{147}$Sm–$^{143}$Nd ages for 60025 are concordant and have a weighted average value of 4,360 ± 3 million years (Myr). This is the first study in which a single FAN has yielded consistent ages from multiple isotopic systems, and strongly suggests that the ages record the time at which the rock crystallized. The Pb–Pb age was determined on sequential dissolutions of a 105.4-mg pyroxene fraction. Four of five dissolution steps of this pyroxene mineral split define a line corresponding to an age of 4,359.2 ± 2.4 Myr (Fig. 1). Several mineral and whole-rock fractions analysed previously[6,14,15] also plot on the 4,359.2 ± 2.4 Myr line but show limited Pb isotopic variation. We interpret the line defined by four dissolution steps of the pyroxene mineral splits to represent a mixture of initial Pb and radiogenic Pb that evolved in a closed system. We note that the Px-2 (L3) fraction has a $^{204}$Pb/$^{206}$Pb ratio of 0.000774, indicating that it is dominated by radiogenic Pb, and consequently defines the *y* intercept on Fig. 1. This is by far the most radiogenic Pb measured in 60025 so far. For example, in a previous detailed study of 60025, the most radiogenic Pb reported was $^{204}$Pb/$^{206}$Pb = 0.0151 (ref. 6). The radiogenic $^{207}$Pb/$^{206}$Pb ratio defined by the *y* intercept is thus well constrained and corresponds to the age of last closure of this system.

The washing and dissolution procedures used to define the Pb–Pb isochron were developed using small (about 10 mg) fractions of pyroxene and plagioclase as test samples. Although the Pb isotopic compositions of these fractions were analysed to assess the efficiency of the procedure, relatively large blank corrections precluded their use in defining the age of 60025. The isotopic compositions of the washes demonstrate that most fall below the isochron towards modern terrestrial Pb. The Px-2 (L2) 1 M hydrofluoric acid (HF) acid fraction also lies towards terrestrial Pb on Fig. 1. Sequential washing and dissolution of primitive chondritic meteorites demonstrates that the first HF digestion step often liberates a large amount of terrestrial Pb contamination[16]. This probably reflects dissolution of small particles, such as dust, in the weak HF acid, which were not removed in the washing steps. It is therefore not surprising that the Px-2 (L2) fraction lies off the isochron towards terrestrial contamination, which justifies removing this fraction from the age calculation. However, including the Px-2 (L2) fraction in the regression only changes the calculated age by about 10 Myr. A few fractions fall above the isochron as well, and are interpreted to have a well known, if poorly understood, lunar Pb contaminant[14,15,17] (see Supplementary Information).

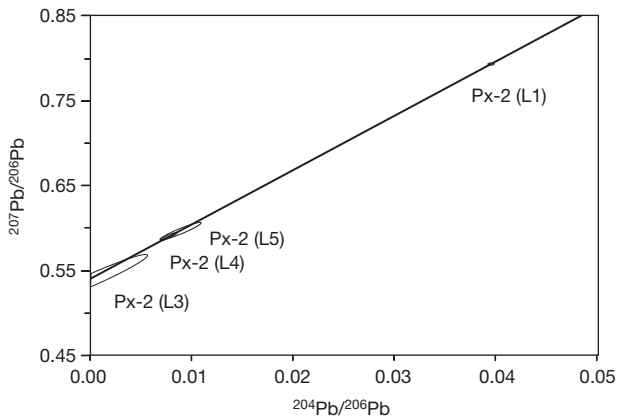



**Figure 1 | Pb–Pb isochron diagram.** For FAN 60025, an age of 4,359.2 ± 2.4 Myr is defined by sequential dissolutions of a 105.4-mg split of the pyroxene mineral fraction. Uncertainties are ±2$\sigma$ mean of population of mass spectrometry ratios plus 50% uncertainty associated with a blank. The mean squared weighted deviation (MSWD) is 1.6.

[1]Chemical Sciences Division, Lawrence Livermore National Laboratory, 7000 East Avenue L-231, Livermore, California 94550, USA. [2]Centre for Star and Planet Formation, University of Copenhagen, Øster Voldgade 5-7 Copenhagen, Denmark. [3]Clermont Université, Université Blaise Pascal, Laboratoire Magmas et Volcans, UMR CNRS 6524, 5 rue Kessler, 63038 Clermont-Ferrand, France. [4]Department of Terrestrial Magnetism, Carnegie Institution, 5241 Broad Branch Road, Northwest, Washington DC 20015-1305, USA..



©2011 Macmillan Publishers Limited. All rights reserved

The Sm–Nd isotopic data define a $^{147}Sm$–$^{143}Nd$ isochron age of 4,367 ± 11 Myr (Fig. 2a) that is 73 Myr younger than that determined previously for 60025 (ref. 5). Recalculating the data of ref. 5 using the IsoPlot program (used to calculate the ages reported here) demonstrates that the two ages differ by only 22 Myr. We note that the much larger sample sizes used here allow significantly higher precision as a result of running Nd as a metal instead of an oxide, as well as minimizing concerns about blank correction (see Supplementary Information). These factors, plus the concordance of Sm–Nd and Pb–Pb ages, suggest that the new result improves on the accuracy of earlier work. In addition, we have produced the first lunar $^{146}Sm$–$^{142}Nd$ internal isochron, yielding an age of $4{,}318^{+30}_{-38}$ Myr (Fig. 2b). Although this age is 41–49 Myr younger than the $^{147}Sm$–$^{143}Nd$ and Pb–Pb ages, its relatively large uncertainty makes it differ from the other ages by only 8–9 Myr. The minor discordance between chronometers probably reflects the small range in measured $^{142}Nd/^{144}Nd$ (48 parts per million) and a slight underestimation of the analytical uncertainty on the measured $^{147}Sm/^{144}Nd$ and/or $^{142}Nd/^{144}Nd$ ratios. This is illustrated in Fig. 2b by the position of the 4,359 Pb–Pb reference line. Regardless of the cause of the very small discrepancy between Sm–Nd ages, the extremely limited variation in $^{142}Nd/^{144}Nd$ between the mineral separates is consistent with the young age determined for 60025.

If 60025 represents a flotation cumulate of the lunar magma ocean, its young crystallization age requires the solidification of the magma ocean, and formation of the Moon, to have occurred later than most previous estimates[18–20]. A young Moon is supported by the observation that almost all lunar crustal rocks[18,19], as well as two FANs[8,9], are the same age or younger than 60025. Although model-dependent, $^{142}Nd/^{144}Nd$ ages for the formation of mare basalt source regions are also young[13,21–23], ranging from $4{,}313^{+25}_{-30}$ Myr to $4{,}353^{+21}_{-23}$ Myr, and are concordant with the $^{143}Nd/^{144}Nd$ model age[24] of 4,360 ± 60 Myr for the last solidification products of the magma ocean (KREEP: potassium, rare-earth elements and phosphate-rich cumulates). Recent $^{182}Hf$–$^{182}W$ measurements have shown no radiogenic input to $^{182}W$ in lunar samples and constrain differentiation of the lunar interior to have occurred after 4.5 billion years (Gyr) ago[25]. The new data for 60025 are thus consistent with formation of the Moon up to about 200 Myr after that of the Solar System.

The oldest terrestrial samples are zircons found in metasedimentary deposits from Jack Hills, Australia, where a single zircon has yielded an age of 4,404 ± 8 Myr (ref. 26). It has REE abundances and an oxygen isotopic composition that indicates it formed from an evolved granitic magma and represents a fragment of continental crust. Taken at face value, the ages of this zircon and 60025 require differentiation on the Moon to occur at least 30 Myr after the differentiation of Earth. This implies that the Moon accreted relatively slowly after the giant impact, or that the Moon retained enough heat to delay cumulate formation.

Alternatively, the young age of 60025 might indicate that it is not a magma ocean product, but rather was produced by a more recent melting event. It should be noted that Sm–Nd ages older than 60025 (Fig. 3) have been reported for three FANs, Y-86032 (4.43 ± 0.03 Gyr; ref. 27), 67075 (4.47 ± 0.07 Gyr; ref. 28) and 67016c (4.53 ± 0.12 Gyr; ref. 10), and one Mg suite sample 15445,17 (4.46 ± 0.07 Gyr; ref. 29). Similarly the oldest lunar zircon, dated at 4,417 ± 6 Myr (ref. 20), comes from the impact melt-breccia 72215 and presumably dates the time of crystallization of a KREEP-rich alkali-suite sample. Finally, some KREEP model ages[18,19,24] are older than 60025, averaging 4,456 ± 65 Myr. Therefore, if 60025 is a flotation cumulate of a magma ocean, these old ages must be in error.

The overlap between the $^{146}Sm$–$^{142}Nd$ isotopic systematics of 60025 and those inferred for mare basalt source regions[13,21–23] (Fig. 2b) places additional constraints on FAN petrogenesis. It suggests that the mare basalt sources and FANs formed with very similar initial Nd isotopic compositions at roughly the same time. Admittedly this overlap is not perfect, but is nonetheless remarkable given that the basalt Nd isotopic measurements were completed under different analytical conditions and variably corrected for thermal neutron capture, and the $^{147}Sm/^{144}Nd$ ratios were calculated using different geochemical models. Geological

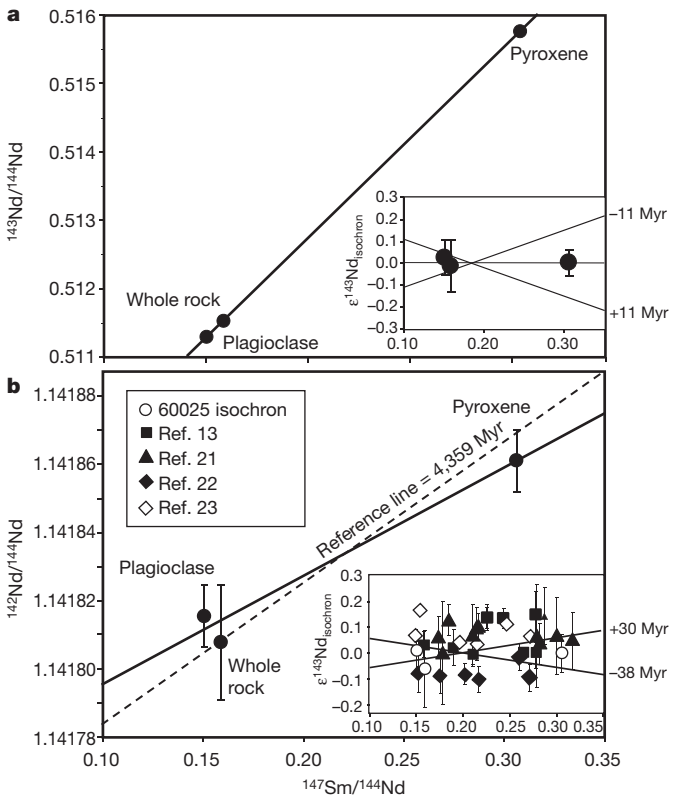



**Figure 2 | Sm–Nd isochron diagrams.** For FAN 60025, plagioclase, pyroxene and whole rocks define a $^{147}Sm$–$^{143}Nd$ age of 4,367 ± 11 Myr (**a**) and a $^{146}Sm$–$^{142}Nd$ age of $4{,}318^{+30}_{-38}$ Myr (**b**). **a**, $\varepsilon Nd = -0.24 \pm 0.09$ (MSWD = 0.40). The inset to **a** represents the deviation of individual points from the isochron in epsilon units. **b**, $\varepsilon^{142}Nd^{CHUR} = +0.10 \pm 0.04$, where CHUR is chondritic uniform reservoir (MSWD = 0.84). $^{146}Sm/^{144}Nd = 0.0085$ at 4,568 Myr ago. The inset to **b** represents deviations of lunar whole rocks analysed in refs 13 and 21–23 from the isochron. Uncertainties are $\pm 2\sigma$ mean of the population of mass spectrometry ratios.

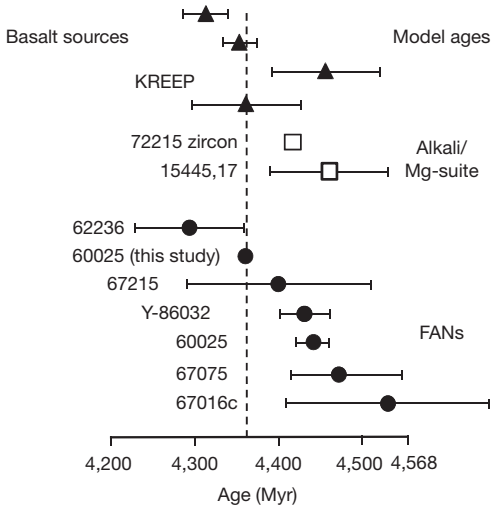



**Figure 3 | Summary of lunar ages.** The diagram illustrates the ages of the oldest lunar samples and model ages of cumulate source regions. Symbols refer to ages reported on individual samples. Filled circles represent FANs, open squares represent Alkali and Mg-suite samples, and filled triangles represent model ages. Error bars are uncertainties reported for individual age determinations. Ages to the right of the dashed line are inconsistent with 60025 representing a flotation cumulate of a lunar magma ocean. Data are referenced in the text and the Supplementary Information.



©2011 Macmillan Publishers Limited. All rights reserved

scenarios for the formation of 60025 must therefore include the formation of the low-Ti (pyroxene-rich), high-Ti (ilmenite-rich) and KREEP-rich source regions of the mare basalts. The success of the magma ocean model of lunar differentiation stems largely from its ability to do just this: account for the petrogenesis of all types of lunar source regions by a single mechanism (see, for example, ref. 4). Thus, if 60025 is not a magma ocean cumulate, then another mechanism to re-equilibrate $^{146}$Sm–$^{142}$Nd isotopic systematics of mafic and plagioclase-rich cumulates is required.

One possible mechanism to equilibrate ancient lunar source regions without invoking magma ocean crystallization is for early-formed cumulates to have overturned as a result of density instability of the cumulate pile (for example, ref. 30). In this scenario of serial magmatism, young FANs represent solidification products of plagioclase-rich melts produced during overturn, whereas mare basalts represent later, more mafic, melts of the same cumulate sources. The advantage of this scenario is that it does not require dismissing the handful of older ages determined on lunar samples and it can potentially account for the variable (if somewhat suspect) ages and widespread distribution of FANs on the lunar surface. The disadvantage is that the lunar magma ocean theory is strongly founded upon the petrological, geochemical and isotopic characteristics of FANs such as 60025. Therefore, if 60025 is not a product of the magma ocean, petrologically similar rocks from the FAN suite cannot be used to characterize its initial solidification. Furthermore, the potential for planetary differentiation by magma ocean solidification on all rocky bodies is weakened if the very rocks that led to the development of the magma ocean theory are themselves not its byproduct.

## METHODS SUMMARY

Thorough examination of the Apollo sample collection at the Johnson Space Center yielded a 1.88-g pyroxene-rich clast from FAN 60025. It contained 25% mafic minerals (mostly pyroxene) that were separated from plagioclase using a Frantz magnetic separator and hand-picking. Ten milligrams of pyroxene and plagioclase were used to develop washing and sequential dissolution procedures designed to remove Pb contamination from the splits. The 10-mg mineral splits were washed in distilled water, ethanol and acetone, and then in weak hydrobromic acid (HBr). They were then sequentially dissolved using a variety of progressively stronger acids. The effectiveness of washing and sequential dissolution to remove Pb contamination was assessed by determining the Pb isotopic composition of individual washes and dissolutions. These analyses demonstrated that only the pyroxene fraction was suitable for age dating. The washing/sequential dissolution procedure developed for the 10-mg pyroxene split was applied to a larger 105.4-mg pyroxene split. The pyroxene was digested in HF and hydrochloric acid (HCl), spiked with a mixed $^{202}$Pb–$^{205}$Pb tracer, and Pb was purified using HBr-$HNO_3$ based chemistry. Lead was run on a Thermo-Fisher Triton by peak hopping in a secondary electron multiplier. Samarium and neodymium were measured on the remaining pyroxene and plagioclase fractions as well as an additional whole-rock fraction. The samples were digested in HF and nitric acid ($HNO_3$) and spiked with a $^{149}$Sm–$^{150}$Nd tracer prepared at Lawrence Livermore National Laboratory and having 99.988% $^{150}$Nd. REEs were separated using HCl and methalactic acid. Total Sm and Nd procedural blanks were 8 pg and 40 pg respectively. The isotopic compositions were determined with a Thermo-Fisher Triton at the Department of Terrestrial Magnetism using double Re filaments. Neodymium data were collected dynamically using 8-s integrations.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 2 March; accepted 20 June 2011.**
**Published online 17 August 2011.**

1. Lapen, T. J. *et al.* A younger age for ALH84001 and its geochemical link to shergottite sources in Mars. *Science* **328,** 347–351 (2010).
2. Wood, J. A., Dickey, J. S., Marvin, U. B. & Powell, B. N. Lunar anorthosites and a geophysical model of the moon. *Proc. Lunar Planet. Sci Conf.* **I,** 965 (1970).
3. Smith, J. A. *et al.* Petrologic history of the Moon inferred from petrography, mineralogy, and petrogenesis of Apollo 11 rocks. *Proc. Lunar Planet. Sci Conf.* **I**, 1, 149 (1970).
4. Snyder, G. A., Taylor, L. A. & Neal, C. R. A chemical model for generating the source of mare basalts: combined equilibrium and fractional crystallization of the lunar magmasphere. *Geochim. Cosmochim. Acta* **56,** 3809–3823 (1992).
5. Carlson, R. W. & Lugmair, G. W. The age of ferroan anorthosite 60025: oldest crust on a young Moon? *Earth Planet. Sci. Lett.* **90,** 119–130 (1988).
6. Hanan, B. B. & Tilton, G. R. 60025—relict of primitive lunar crust. *Earth Planet. Sci. Lett.* **84,** 15–21 (1987).
7. Norman, M. D., Borg, L. E., Nyquist, L. E. & Bogard, D. D. Chronology, geochemistry, and petrology of a ferroan noritic anorthosite clast from Descartes breccia 67215: clues to the age, origin, structure, and impact history of the lunar crust. *Meteorit. Planet. Sci.* **38,** 645–661 (2003).
8. Borg, L. E. *et al.* Isotopic studies of ferroan anorthosite 62236: a young lunar crustal rock from a light rare-earth element-depleted source. *Geochim. Cosmochim. Acta* **63,** 2679–2691 (1999).
9. Alibert, C., Norman, M. D. & McCulloch, M. T. An ancient age for a ferroan anorthosite clast from lunar breccia 67016. *Geochim. Cosmochim. Acta* **58,** 2921–2926 (1994).
10. Shirley, D. N. A partially molten magma ocean model. *Proc. Lunar Planet. Sci Conf.* **XII,** A519–A527 (1983).
11. Warren, P. H. & Wasson, J. T. The origin of KREEP. *Rev. Geophys. Space. Phys.* **17,** 73–88 (1977).
12. James, O. B., Lindstrom, M. M. & McGee, J. J. Lunar ferroan anorthosites 60025—petrology and chemistry of mafic lithologies. *Proc. Lunar Planet. Sci. Conf.* **XXI,** 63–87 (1991).
13. Boyet, M. & Carlson, R. W. A highly depleted Moon or a non-magma origin for the lunar crust? *Earth Planet. Sci. Lett.* **262,** 505–516 (2007).
14. Tera, F. & Wasserburg, G. J. U-Th-Pb systematics in the lunar highland samples from the Luna 20 and Apollo 16 missions. *Earth Planet. Sci. Lett.* **17,** 36–51 (1972).
15. Nunes, P. D., Knight, R. J., Unruh, D. M. & Tatsumoto, M. The primitive nature of the lunar crust and the problem of initial Pb isotopic compositions of lunar rocks: a Rb-Sr and U-Th-Pb study of Apollo 16 samples. *Lunar Planet. Sci. Conf.* **V,** 559–561 (1974).
16. Connelly, J. N. & Bizzarro, M. Pb/Pb dating of chondrules from CV chondrites by progressive dissolution. *Chem. Geol.* **259,** 143–151 (2009).
17. Tera, F. & Wasserburg, G. J. U-Th-Pb systematics of lunar rocks and inferences about lunar evolution. *Proc. Lunar Sci. Conf.* **V,** 1571–1599 (1974).
18. Nyquist, L. E. & Shih, C.-Yu. The isotopic record of lunar volcanism. *Geochim. Cosmochim. Acta* **56,** 2213–2234 (1992).
19. Edmunson, J., Borg, L. E., Nyquist, L. E. & Asmeron, Y. A combined Sm-Nd, Rb-Sr, and U-Pb isotopic study of Mg-suite norite 78238: further evidence for early differentiation of the Moon. *Geochim. Cosmochim. Acta* **73,** 514–527 (2009).
20. Nemchin, A. *et al.* Timing of crystallization of the lunar magma ocean constrained by the oldest zircon. *Nature Geosci.* **25,** 133–136 (2009).
21. Nyquist, L. E. *et al.* $^{146}$Sm-$^{142}$Nd formation interval for the lunar mantle. *Geochim. Cosmochim. Acta* **59,** 2817–2837 (1995).
22. Rankenburg, K., Brandon, A. D. & Neal, C. R. Neodymium isotope evidence for a chondritic composition of the Moon. *Science* **312,** 1369–1372 (2006).
23. Brandon, A. D. *et al.* Re-evaluating $^{142}$Nd/$^{144}$Nd in lunar mare basalts with implications for early evolution and bulk Sm/Nd of the Moon. *Geochim. Cosmochim. Acta* **73,** 6421–6445 (2009).
24. Carlson, R. W. & Lugmair, G. W. The Sm-Nd history of KREEP. *Proc. Lunar Planet. Sci Conf.* **IX,** 689–704 (1978).
25. Touboul, M., Kleine, T., Bourdon, B., Palme, H. & Wieler, R. Late formation and prolonged differentiation of the Moon inferred from W isotopes in lunar metals. *Nature* **450,** 1206–1209 (2007).
26. Wilde, S. A., Valley, J. W., Peck, W. H. & Graham, C. M. Evidence for detrital zircons for the existance of continental crust and oceans on Earth 4.4 Gyr ago. *Science* **409,** 175–178 (2001).
27. Nyquist, L. *et al.* Feldspathic clasts in Yamato-86032: remnants of the lunar crust with implications for its formation and impact history. *Geochim. Cosmochim. Acta* **70,** 5990–6015 (2006).
28. Nyquist, L. E. *et al.* Lunar crustal history recorded in lunar anorthosites. *Lunar Planet. Sci Conf.* **XXXXI,** 1383 (2010).
29. Shih, C.-Y. *et al.* Age of pristine noritic clasts from lunar breccias 15445 and 15455. *Geochim. Cosmochim. Acta* **57,** 915 (1993).
30. Hess, P. C. & Parmentier, E. M. A model for the thermal and chemical evolution of the Moon's interior; implications for the onset of mare volcanism. *Earth Planet. Sci. Lett.* **134,** 501–514 (1995).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** This work was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under contract number DE-AC52-07NA27344. The portion of the work performed at Lawrence Livermore National Laboratory and the Department of Terrestrial Magnetism were supported by NASA Cosmochemistry grants NNH08ZDA001N and NNX08AH65G, respectively. The Centre for Star and Planet Formation is funded by the Danish National Research Foundation and the University of Copenhagen's programme of excellence. We appreciate comments by A. Brandon.

**Author Contributions** L.E.B. identified, located and prepared the sample for analysis. J.N.C. performed Pb–Pb measurements. L.E.B., M.B. and R.W.C. completed Sm–Nd measurements. All authors contributed to interpretation of data and preparation of the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to L.E.B. (borg5@llnl.gov).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Sample preparation.** L.E.B. visited the lunar curation facility at the Johnson Space Center (JSC) in October 2008 and examined several pieces of FAN 60025 searching for a clast that was enriched in mafic minerals. The sample allocated at the JSC weighed 1.88 g and contained approximately 25% mafic minerals consisting of olivine and pyroxene intergrown with large grains of plagioclase. The sample was crushed into coarse fragments, creating some fine powder. Several large fragments were set aside. Mafic and plagioclase fragments were high-graded into two fractions using tweezers. There remained a third fraction that was a 50/50 mixture of mafic minerals and plagioclase, but was too fine-grained to be high-graded using the tweezers. The coarse mafic fraction and the 50/50 mixture were crushed in a sapphire mortar and pestle and sieved yielding fractions of 75–200, 200–325 and <325 mesh. Mineral separations were completed on these fractions using a Frantz isobarrier separator and handpicking under dry conditions to minimize Pb contamination during processing. Pyroxene was also separated from the <325 mesh split of the coarse-grained mafic fraction but was too fine-grained to be handpicked. Visual inspection of the mafic mineral and plagioclase fractions indicated that they were >99% pure.

**Pb procedures.** The Pb isotopic data were collected in three analytical sessions. The first two sessions were preliminary and used to develop washing and sequential multi-step dissolution procedures to remove contamination from 10-mg splits of plagioclase and pyroxene mineral fractions. The Pb isotopic compositions of individual washes were analysed throughout the washing procedure to determine how much Pb contamination was removed during each washing step, and to assess the next step in the procedure. Once we were convinced that significant quantities of contamination had been removed, sequential digestion of the samples was begun. Preliminary Pb isotopic analysis of wash and sequential digestion fractions of the 10-mg plagioclase split demonstrated that not enough Pb contamination was removed to warrant additional analysis of plagioclase in 60025. The third and final analytical session therefore focused on a 105.4-mg split of the pyroxene fraction. Washing and sequential digestion procedures used in the third session were similar, but not identical, to those used in the preliminary session on the 10-mg pyroxene split. We note that the sequential digestions of the 105.4-mg pyroxene fraction were the only dissolutions designed to define the age of the sample. Nevertheless, the isotopic compositions of all of the preliminary washes and digestions of the plagioclase and both pyroxene splits are reported in the Supplementary Information.

The fractions for all sessions were pre-washed in five cycles of distilled water, ethanol and acetone before progressive weak-acid washing. The Pb isotopic composition of these washes was not determined. Next, the 10-mg splits were washed in water followed by six washes in very weak HBr. The final wash was in 0.5 M HBr. After washing, the 10-mg plagioclase and pyroxene splits were sequentially dissolved in a multi-step procedure that used progressively stronger acids. Despite the small amount of Pb in the final dissolution steps of the 10-mg pyroxene split, the analyses of the washes and sequential dissolutions of this fraction showed two encouraging features for geochronology: (1) as cleaning progressed, the Pb from the residual material became more radiogenic, and (2) the analyses of the sequential dissolutions defined an approximately linear trend despite a significant correction for Pb blank on most analyses.

From this information it was clear that the pyroxene required extensive cleaning in water, acetone, ethanol and weak acids, with a final cleaning step using 0.5 M HBr before starting the multi-step dissolution. The dissolution procedure successively used 6 M HCl, 1 M HF, 7 M HF, 7 M $HNO_3$, and 28 M HF + 14 M $HNO_3$ to produce the Px-2 (L1)–(L5) fractions, respectively. Lead was separated from matrix elements by passing the samples twice through a HBr–$HNO_3$-based chemical separation procedure using 0.055 ml of Eichrom anion resin[16]. The purified Pb was analysed on a Thermo-Fisher Triton thermal ionization mass spectrometer equipped with nine Faraday detectors and one axial secondary-electron-multiplier ion-counting system. Lead was loaded onto previously outgassed zone-refined Re filaments with silica gel made from silicic acid following ref. 31. All analyses were made by sequentially peak jumping the ion beams into the central secondary-electron-multiplier ion-counting system. Samples were spiked with a $^{202}Pb$–$^{205}Pb$ double spike to allow for an internal correction for instrument mass fractionation. Samples are corrected for a Pb blank added during the chemical separation according to replicate blanks run during the same session. Lead blanks from the chemistry during this work ranged from 0.25 to 0.50 pg. Samples are also corrected for a 0.2-pg Pb loading blank. The data with 2$\sigma$ errors are presented in the Supplementary Information.

**Sm–Nd procedures.** Samples for Sm–Nd analyses were derived from the aliquots of the plagioclase and whole-rock fractions used for Pb–Pb analyses, plus an additional whole-rock fraction. The samples were dissolved in 3:1 mixtures of concentrated HF:$HNO_3$ in sealed Savillex beakers on a hot plate at 90 °C. Before dissolution, the samples were spiked with $^{149}Sm$ and $^{150}Nd$, for concentration determination. A new mixed spike consisting of 99.988% purity $^{150}Nd$ and 97.7% $^{149}Sm$ was prepared at Lawrence Livermore National Laboratory for this work and calibrated against the AMES metal Sm–Nd mixed standard described in ref. 32. After three days on the hot plate, the samples were evaporated to dryness, treated twice with concentrated $HNO_3$, drying in between treatments, and then dissolved in 6 M HCl. The volume of HCl was increased until a clear solution was obtained. Once complete dissolution was achieved, the samples were dried and then redissolved in 5 ml of 1 M HCl + 0.1 M HF. The samples were centrifuged to remove any undissolved material. The clear liquid was loaded on 20-cm-long, 1-cm-diameter quartz columns filled with AG50W-X8 resin. High-field-strength elements were eluted in the sample loading solution and an additional 5 ml of 1 M HCl/0.1 M HF. Major elements and large ion lithophile trace elements were then eluted in 2.5 M HCl with heavy-REE and light-REE splits eluted in 4 M HCl. The light-REE split was processed further to separate Sm and Nd using the procedure described most recently in ref. 33. The column procedure used here has a total blank of 12 pg Nd and 2.3 pg Sm. The sample data were corrected using blanks of 40 pg Nd and 8 pg Sm to account for the volume of acid used for dissolution and because each sample was split between two columns for the first step of the chemical separation. The smallest Sm–Nd amounts analysed here were the 89 ng of Nd obtained from the mafic (olivine + pyroxene) mineral split and the 40 ng of Sm extracted from the plagioclase. The blanks listed above thus constitute less than 0.05% of the analysed Sm and Nd and are negligible. Concentration data are presented in the Supplementary Information.

Neodymium was loaded in 3 M HCl onto a Re filament and analysed as $Nd^+$ using a second Re filament for ionization. Isotope ratios were measured with the Department of Terrestrial Magnetism's Thermo-Fisher Triton thermal ionization mass spectrometer using a two-mass-step procedure with $^{145}Nd$ and then $^{143}Nd$ in the axial Faraday detector. This procedure calculates $^{142}Nd/^{144}Nd$ dynamically to eliminate Faraday detector biases. All other Nd-isotope ratios are measured statically along with potential interfering species Ce and Sm. Each step integrates the signal for 8 s, leading to 16 s of signal integration for each ratio except for $^{142}Nd/^{144}Nd$ and $^{150}Nd/^{144}Nd$ that have 8 s of signal integration per ratio. For the plagioclase, 480 ratios were obtained at an average $^{142}Nd = 1.2 \times 10^{-11}$ A. The pyroxene produced 420 ratios at an average $^{142}Nd = 1.6 \times 10^{-11}$ A. The whole-rock analysis lasted only for 130 ratios at an average $^{142}Nd = 8.5 \times 10^{-12}$ A. Errors reported for the sample measurements are the 2$\sigma$-mean from the individual mass spectrometer runs. All data are corrected ratio by ratio during mass spectrometry for Ce interference (using $^{142}Ce/^{140}Ce = 0.125$) and Sm interference (using the measured Sm isotopic composition of each sample). Six 540 ratio runs of the JNdi Nd standard at signal sizes of $2–5 \times 10^{-11}$ A were run in the same barrel as the sample analyses. The results of these standard runs are listed in the Supplementary Information along with the isotopic compositions measured for each sample. Because the samples were total-spiked, the $^{150}Nd/^{144}Nd$ and $^{149}Sm/^{152}Sm$ ratios are considerably higher than standard Nd and Sm. All Nd ratios except for $^{146}Nd/^{144}Nd$ are corrected for mass fractionation to $^{146}Nd/^{144}Nd = 0.7219$, using exponential mass dependency. All but $^{146}Nd/^{144}Nd$ and $^{150}Nd/^{144}Nd$ are corrected for the small spike contribution to these isotopes from the highly enriched $^{150}Nd$ spike. The $^{146}Nd/^{144}Nd$ ratio reported in the Supplementary Information is the measured value before fractionation and spike correction.

Samarium was loaded in 2 M $HNO_3$ onto a Re filament and analysed as $Sm^+$ using a second Re filament for ionization. All Sm isotopes along with potential interference from Nd ($^{146}Nd$) were measured statically for 180 ratios of 8-s integration each. Corrections for Nd interference on Sm were made using the measured Nd isotopic composition of each sample. Sm concentrations were calculated assuming normal Sm as previous measurements of unspiked Sm from 60025 show no resolvable deficit in $^{149}Sm$ that could be attributable to neutron capture[13,24]. Both the whole-rock and mafic fractions measured here have spike-corrected $^{150}Sm/^{152}Sm$ ratios only marginally higher than normal Sm, in agreement with previous unspiked Sm measurements[13,24] in different samples of 60025. The Sm isotopic compositions are corrected for fractionation to $^{147}Sm/^{152}Sm = 0.56081$ with the exception of the $^{147}Sm/^{152}Sm$ ratio, which is reported as measured in the Supplementary Information. All ratios except $^{149}Sm/^{152}Sm$ and $^{147}Sm/^{152}Sm$ reported in the Supplementary Information have also been corrected for the contribution from the $^{149}Sm$ spike.


31. Gerstenberger, H. & Haase, G. A highly effective emitter substance for mass spectrometric Pb isotope ratio determinations. *Chem. Geol.* **136,** 309–312 (1997).
32. Boyet, M. & Carlson, R. W. $^{142}Nd$ evidence for early (>4.53 Ga) global differentiation of the silicate Earth. *Science* **309,** 576–581 (2005).
33. Boyet, M., Carlson, R. W. & Horan, M. Old Sm-Nd ages for cumulate eucrites and redetermination of the solar system initial $^{146}Sm/^{144}Sm$ ratio. *Earth Planet. Sci. Lett.* **291,** 172–181 (2010).


©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10329

# TSLP promotes interleukin-3-independent basophil haematopoiesis and type 2 inflammation

Mark C. Siracusa[1,2], Steven A. Saenz[1,2], David A. Hill[1,2], Brian S. Kim[1,2], Mark B. Headley[3], Travis A. Doering[1,2], E. John Wherry[1,2], Heidi K. Jessup[4], Lori A. Siegel[4], Taku Kambayashi[5], Emily C. Dudek[6], Masato Kubo[7,8], Antonella Cianferoni[6], Jonathan M. Spergel[6], Steven F. Ziegler[3], Michael R. Comeau[4] & David Artis[1,2,9]

**$CD4^+$ T-helper type 2 ($T_H2$) cells, characterized by their expression of interleukin (IL)-4, IL-5, IL-9 and IL-13, are required for immunity to helminth parasites[1] and promote the pathological inflammation associated with asthma and allergic diseases[2]. Polymorphisms in the gene encoding the cytokine thymic stromal lymphopoietin (TSLP) are associated with the development of multiple allergic disorders in humans, indicating that TSLP is a critical regulator of $T_H2$ cytokine-associated inflammatory diseases[3–6]. In support of genetic analyses, exaggerated TSLP production is associated with asthma, atopic dermatitis and food allergies in patients, and studies in murine systems demonstrated that TSLP promotes $T_H2$ cytokine-mediated immunity and inflammation[5,7–12]. However, the mechanisms through which TSLP induces $T_H2$ cytokine responses remain poorly defined. Here we demonstrate that TSLP promotes systemic basophilia, that disruption of TSLP–TSLPR interactions results in defective basophil responses, and that TSLPR-sufficient basophils can restore $T_H2$-cell-dependent immunity *in vivo*. TSLP acted directly on bone-marrow-resident progenitors to promote basophil responses selectively. Critically, TSLP could elicit basophil responses in both IL-3–IL-3R-sufficient and -deficient environments, and genome-wide transcriptional profiling and functional analyses identified heterogeneity between TSLP-elicited versus IL-3-elicited basophils. Furthermore, activated human basophils expressed TSLPR, and basophils isolated from eosinophilic oesophagitis patients were distinct from classical basophils. Collectively, these studies identify previously unrecognized heterogeneity within the basophil cell lineage and indicate that expression of TSLP may influence susceptibility to multiple allergic diseases by regulating basophil haematopoiesis and eliciting a population of functionally distinct basophils that promote $T_H2$ cytokine-mediated inflammation.**

Previous *in vitro* studies suggested that TSLP could promote $T_H2$ cytokine-mediated inflammation by influencing dendritic cell, mast cell and lymphocyte populations[13]. However, the IL-4-expressing cell populations that TSLP targets *in vivo* remain poorly defined. Following administration of recombinant TSLP (rTSLP) to IL-4/eGFP (enhanced green fluorescent protein) reporter mice, an increase in the frequency and total number of basophils was observed in the spleen (Fig. 1a, b). Mice receiving hydrodynamic tail vein injections of cDNA plasmid encoding TSLP (TSLP cDNA)[14] also exhibited significant increases in the frequency and total number of basophils in the spleen (Fig. 1c, d), blood, lung and bone marrow (Supplementary Fig. 1a–f). Basophils, along with dendritic cells, are capable of contributing to the induction of optimal $T_H2$ cytokine responses[15–20]. Consistent with this, exaggerated basophil responses in TSLP-cDNA-treated mice were associated with increased systemic production of $T_H2$ cytokines

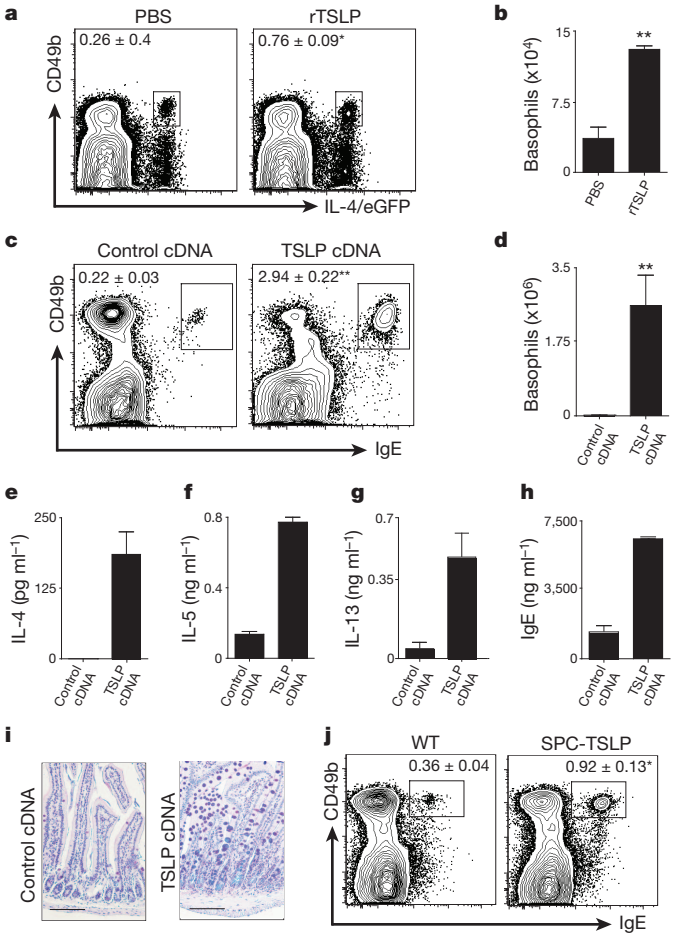



**Figure 1 | TSLP promotes peripheral basophilia.** **a**, **b**, IL-4/eGFP reporter mice were treated with PBS or rTSLP and splenic basophils were identified (**a**) and quantified (**b**). **c**, **d**, Wild-type mice were treated with control cDNA or TSLP-encoding (TSLP cDNA) plasmid and splenic basophils were identified (**c**) and quantified (**d**). **e**–**h**, Serum IL-4 (**e**), IL-5 (**f**), IL-13 (**g**) and total IgE (**h**) were quantified. **i**, Histological analysis of small intestine. Scale bars, 100 μm. **j**, Splenic basophils from control mice (WT), or SPC-TSLP mice were identified. Flow cytometry plots are gated on live NBNT cells. Results are representative of at least three (**a**–**d** and **j**–**i**) or two (**e**–**i**) separate experiments containing 3–5 (**a**–**d**), 5 (**e**–**i**) or 6–12 (**j**) mice per group. Results are shown as means ± s.e.m. Statistical analyses performed using a two-tailed students *t*-test (*$P < 0.05$), (**$P < 0.01$).

[1]Institute for Immunology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [2]Department of Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [3]Immunology Program, Benaroya Research Institute, Seattle, Washington 98101, and Immunology Department, University of Washington School of Medicine, Seattle, Washington 98195, USA. [4]Inflammation Research, Amgen Inc., Seattle, Washington 98119, USA. [5]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [6]Division of Allergy and Immunology, Children's Hospital of Philadelphia, Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [7]Laboratory for Signal Network, Research Center for Allergy and Immunology, RIKEN Yokohama Institute, Yokohama 230-0045, Japan. [8]Division of Molecular Pathology, Research Institute for Biological Science, Tokyo University of Science, Chiba 278-0022, Japan. [9]Department of Pathobiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.



©2011 Macmillan Publishers Limited. All rights reserved

(Fig. 1e–g), elevated IgE levels (Fig. 1h) and the development of $T_H2$ cytokine-associated intestinal inflammation (Fig. 1i). In addition to reported expression by mast cells and basophils, epithelial cells at barrier surfaces are potent sources of TSLP[13]. Therefore, we tested whether epithelial-cell-derived TSLP could influence peripheral basophil populations. Mice that constitutively express a *Tslp* transgene under the surfactant protein C promoter (SPC-TSLP) in lung epithelial cells[21] exhibited increased frequencies of basophils in the spleen (Fig. 1j).

To test whether TSLP–TSLPR signalling influences infection-induced basophilia *in vivo*, we used a helminth model in which protective immunity is dependent on TSLP–TSLPR signalling[10]. Whereas *Trichuris muris*-infected wild-type mice exhibited increased basophils in the periphery, $Tslpr^{-/-}$ mice failed to exhibit infection-induced basophilia (Supplementary Fig. 2a). Critically, TSLP-elicited basophils isolated from $Tslpr^{+/+}$ mice and adoptively transferred into normally susceptible *T. muris*-infected $Tslpr^{-/-}$ mice were sufficient to restore $T_H2$ cytokine responses (Supplementary Fig. 2b,c), IgE production (Supplementary Fig. 2d), goblet cell hyperplasia (Supplementary Fig. 2e, f), secretion of goblet-cell-derived RELM-β (Supplementary Fig. 2g) and significantly reduce worm burdens (Supplementary Fig. 2h). These data demonstrate that basophil-restricted TSLPR expression is sufficient to partially restore protective $T_H2$ cytokine responses *in vivo*.

TSLP was originally identified as a lymphocyte growth factor capable of regulating the differentiation of lymphocytes from haematopoietic progenitor cells[13]. These findings provoked the hypothesis that in addition to influencing peripheral basophil responses, TSLP may regulate basophil haematopoiesis in the bone marrow. To test directly whether TSLP could elicit the population expansion of basophils from bone-marrow-resident cells, we adopted a basophil differentiation assay using IL-4/eGFP reporter mice[22]. When bone marrow was cultured in media alone, 4.25% of the non-B non-T (NBNT) cells expressed IL-4/eGFP (Fig. 2a), but no mast cells or basophils were present (Fig. 2b). As previously reported[22], bone marrow cells cultured

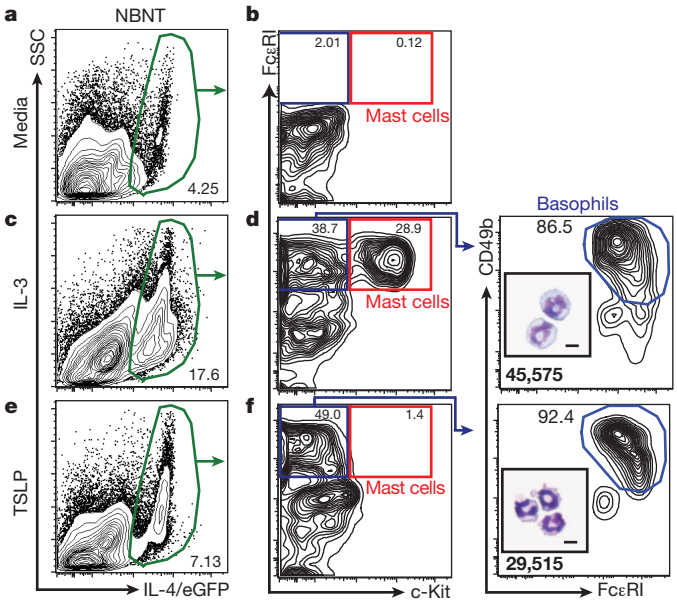



**Figure 2 | TSLP preferentially expands basophil populations from bone-marrow-resident cells.** **a**, **c**, **e**, Bone-marrow-resident cells were taken from IL-4/eGFP reporter mice and cultured in the presence of media (**a**), IL-3 (**c**) or TSLP (**e**), and the frequency of IL-4/eGFP$^+$ NBNT cells was determined. **b**, **d**, **f**, The frequencies of IL-4/eGFP$^+$, NBNT, Siglec-F$^-$ basophils and mast cells were determined. Insets of cytology of IL-4/eGFP$^+$, CD49b$^+$, FcεRI$^+$ and c-Kit$^-$ IL-3-elicited (**d**) or TSLP-elicited (**f**) basophils are shown. Results are representative of at least three separate experiments. Bold numbers in basophil gates represent the mean florescence intensity (MFI) levels of CD49b staining. Scale bar, 10 μm.

in the presence of IL-3 exhibited increased frequencies of IL-4/eGFP$^+$ NBNT cells (17.6%) (Fig. 2c) and increased mast cell and basophil populations (Fig. 2d). In contrast, TSLP-stimulated bone marrow cultures yielded increased frequencies of IL-4/eGFP$^+$ cells (7.13%) (Fig. 2e) and basophils but did not yield mast cell populations (Fig. 2f). Consistent with these data, bone marrow from wild-type mice cultured in the presence of IL-3 exhibited increased frequencies and total numbers of basophils and mast cells, whereas bone marrow cultured in the presence of TSLP alone exhibited increased frequencies and total numbers of basophils but not mast cells (Supplementary Fig. 3a–h). Collectively, these data illustrate that TSLP can elicit the selective population expansion of mature basophils from bone-marrow-resident cells.

Basophil populations develop from precursors (BaP) that are characterized as CD34$^+$c-Kit$^-$FcεRI$^+$NBNT cells[23]. BaPs isolated from bone marrow expressed both chains of the TSLPR complex (IL-7Rα and TSLPR), indicating that they may be TSLP-responsive (Supplementary Fig. 4a). To test whether IL-3 or TSLP could induce basophil maturation, sorted bone-marrow-resident BaPs (Supplementary Fig. 4b) were cultured in the presence of media, IL-3 or TSLP and mature basophil populations were evaluated on day 5 after culture. Greater than 30% of BaPs cultured in the presence of media alone retained their progenitor-like phenotype as measured by expression of CD34 (Supplementary Fig. 4c). In contrast, IL-3-stimulated BaPs were predominantly CD34$^-$ and exhibited elevated CD49b expression, a phenotype consistent with mature basophils (Supplementary Fig. 4c, d). Critically, TSLP-stimulated BaPs were also predominantly CD34$^-$ and exhibited elevated levels of CD49b expression (Supplementary Fig. 4c, d). Consistent with whole bone marrow cultures (Fig. 2d, f), IL-3- or TSLP-treated cell cultures yielded more basophils than media alone controls, and IL-3-elicited basophils were larger in size than TSLP-elicited basophils (Supplementary Fig. 4e, f). To compare the ability of IL-3 or TSLP to enhance basophil survival, mature basophils were sort-purified from the blood and spleen of wild-type mice and cultured in the presence of media, IL-3 or TSLP for 24 h. Whereas most mature basophils died when cultured in media alone, 24.6% of basophils survived in the presence of IL-3 and 13.1% of the basophils survived in the presence of TSLP (Supplementary Fig. 5a). These data demonstrate that both IL-3 and TSLP can promote basophil maturation from bone-marrow-resident precursor cells and promote survival of mature basophils isolated from the periphery.

IL-3 is reported to be essential for optimal basophil activation, population expansion and survival[24]. To test whether TSLP-elicited basophil responses were dependent on IL-3–IL-3R signalling, bone marrow cells from wild-type mice or mice deficient in both IL-3Rβil3 and IL-3Rβc ($Csf2rb2^{-/-}Csf2rb^{-/-}$)[25] were cultured in the presence of media, IL-3 or TSLP. Addition of IL-3 to wild-type bone marrow cultures increased the frequency of basophils compared to media alone controls (Fig. 3a), whereas the addition of IL-3 to $Csf2rb2^{-/-}Csf2rb^{-/-}$ bone marrow cultures had no effect on basophil populations (Fig. 3b). Similar to IL-3 stimulation, addition of TSLP to wild-type bone marrow cultures increased the frequency of basophils (Fig. 3a). Critically, addition of TSLP to $Csf2rb2^{-/-}Csf2rb^{-/-}$ bone marrow cultures also increased the frequency of basophils (Fig. 3b), demonstrating that TSLP can promote the population expansion of basophils from bone marrow cells independently of IL-3–IL-3R signalling. However, basophils were readily identified in the spleen and bone marrow of $Csf2rb2^{-/-}Csf2rb^{-/-} \times Tslpr^{-/-}$ mice, or $Csf2rb2^{-/-}Csf2rb^{-/-}$ mice treated with neutralizing anti-TSLP monoclonal antibody (Supplementary Fig. 5b, c), demonstrating that signalling via IL-3–IL-3R and TSLP–TSLPR is not essential for basophil development. To test whether TSLP could elicit mature basophil responses independently of IL-3–IL-3R signalling *in vivo*, wild-type or $Csf2rb2^{-/-}Csf2rb^{-/-}$ mice were injected with PBS or rTSLP and peripheral basophil responses were examined. TSLP-treated wild-type mice exhibited increased frequencies and total numbers of basophils



©2011 Macmillan Publishers Limited. All rights reserved

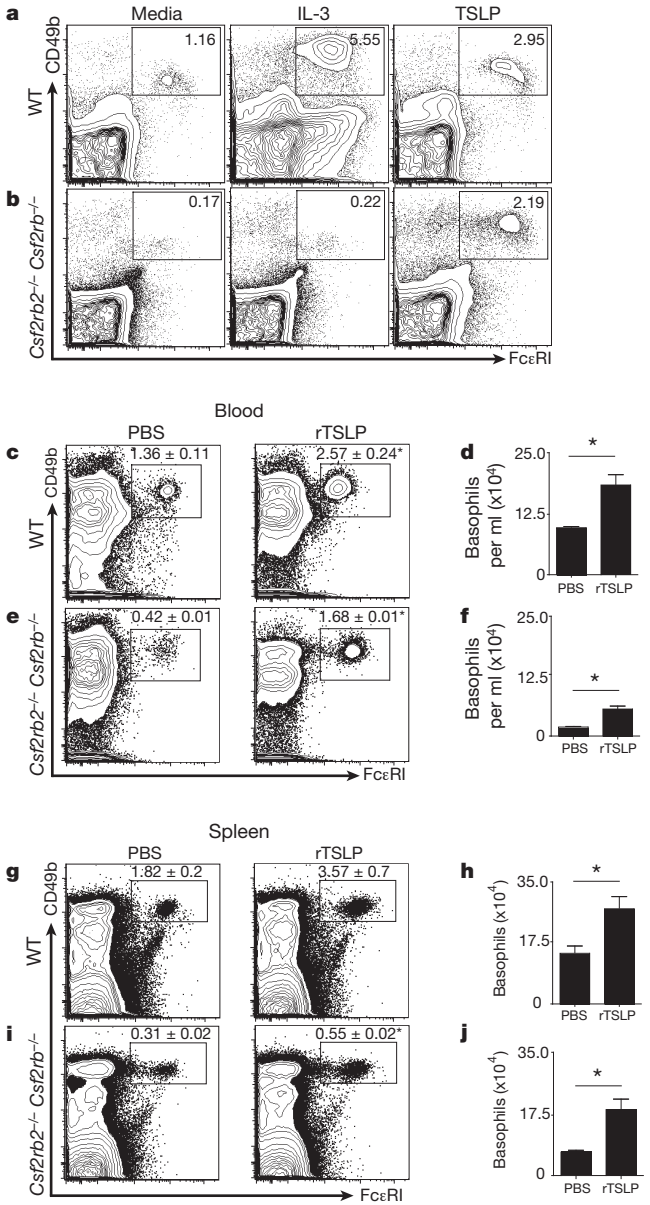



**Figure 3 | TSLP-elicited basophilia is independent of IL-3–IL-3R signalling.** **a**, **b**, Bone marrow cultures from wild-type (WT; **a**) or *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ (**b**) mice. **c**, **e**, **g**, **i**, Wild-type or *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ mice were treated with PBS or rTSLP and blood (**c**, **e**) and splenic basophils (**g**, **i**) were identified. **d**, **f**, **h**, **j**, Total numbers of basophils were quantified. Flow cytometry plots are gated on live NBNT cells. **a**, **b**, Results are representative of at least 3 separate experiments. **c**–**j**, Results are representative of 3 separate experiments (WT + PBS, $n = 7$; *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ + PBS, $n = 7$; WT + TSLP, $n = 7$; *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ + TSLP, $n = 7$). Results are shown as means ± s.e.m. Statistical analysis was performed using a two-tailed students *t*-test (*$P < 0.05$).

in the blood (Fig. 3c, d) and spleen (Fig. 3g, h) compared to PBS treated controls. Critically, rTSLP treatment significantly increased both the frequency and total number of basophils in the blood (Fig. 3e, f) and spleen (Fig. 3i, j) of *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ mice. Collectively, these data illustrate that TSLP promotes peripheral basophilia in both IL-3–IL-3R-sufficient and -deficient environments.

To test if IL-3-elicited versus TSLP-elicited basophil populations were phenotypically distinct, flow cytometric analysis of basophil-associated surface markers was performed. Whereas IL-3-elicited and TSLP-elicited basophils exhibited similar expression of CD200R and CD69, IL-3-elicited basophils exhibited elevated expression of CD11b and CD62L (Fig. 4a). In contrast, TSLP-elicited basophils exhibited elevated expression of CD123, IL-33R (also called T1/ST2) and IL-18Rα (Fig. 4b). Collectively, these data demonstrate that IL-3- versus TSLP-elicited basophils exhibit distinct surface marker expression. Genome-wide transcriptional profiling of IL-3-elicited versus TSLP-elicited basophils also revealed remarkable differences in patterns of gene expression between the two basophil populations. Gene-set enrichment analysis (GSEA)[26] of complete transcriptional profiles revealed that IL-3-elicited basophils reflected gene expression associated with monocyte and dendritic cell maturation, matrix metalloproteinase production and tumour-necrosis factor-α (TNF-α) signalling (Fig. 4c and Supplementary Fig. 6). In contrast, TSLP-elicited basophils exhibited a gene expression profile enriched for linoleic acid metabolism, arachidonic acid metabolism, cell communication and the expression of cell adhesion molecules (Fig. 4c and Supplementary Fig. 6). Real-time PCR analysis also revealed selective differential expression of messenger RNA encoding basophil-associated proteases mast cell protease 2 (*Mcpt2*) and *Mcpt7* (also called *Tpsab1*) between the two populations (Supplementary Fig. 7a). Collectively, these results demonstrate phenotypic heterogeneity between IL-3-elicited versus TSLP-elicited basophils and indicate that either the two basophil populations represent distinct developmental stages of basophils or that they are functionally heterogeneous mature basophil populations.

To test whether phenotypic differences in the two basophil populations were associated with functional heterogeneity, a panel of basophil activation assays was used. IL-3-elicited basophils exhibited similar degranulation and β-hexosaminidase release[27] (a hallmark of basophil effector function[28]) as mast cells in response to IgE-mediated FcεRI crosslinking (Supplementary Fig. 7b). Consistent with distinct effector functions, TSLP-elicited basophils exhibited minimal degranulation (Supplementary Fig. 7b). Surface marker expression on IL-3-elicited versus TSLP-elicited basophils indicated that these two populations may differ in their responsiveness to cytokines (Fig. 4b). To test this directly, IL-3- or TSLP-elicited basophils were sort-purified and stimulated with media alone, IL-3, IL-18 or IL-33, and production of cytokines and chemokines was assessed. Whereas the two basophil populations shared some functions including similar production of IL-1α and CCL11 (Supplementary Fig. 8a, b), TSLP-elicited basophils produced significantly more IL-4 and IL-6 (Supplementary Fig. 8a) and more CCL3, CCL4 and CCL12 (Supplementary Fig. 8b) in response to IL-3 stimulation compared to IL-3-elicited basophils. TSLP-elicited basophils also produced significantly more IL-4, IL-6, TNF-α (Supplementary Fig. 8a), CCL3, CCL4, CCL9, CCL12 (Supplementary Fig. 8b) and CXCL2 (Supplementary Fig. 8c) in response to IL-18 stimulation when compared to IL-3-elicited basophils. Finally, TSLP-elicited basophils produced more IL-4, IL-6 (Supplementary Fig. 8a), CCL3, CCL12 (Supplementary Fig. 8b) and CXCL2 (Supplementary Fig. 8c) in response to IL-33 stimulation than IL-3-elicited basophils. Collectively, these data indicate that IL-3-elicited and TSLP-elicited basophils exhibit distinct responsiveness and functional potential after stimulation with IL-3, IL-18 or IL-33.

To test whether IL-3-independent, TSLP-dependent basophil responses exist *in vivo* following exposure to an inflammatory stimulus, a TSLP-dependent murine model of atopic dermatitis[29] was used. Topical treatment with the vitamin D analogue MC903 resulted in increased levels of TSLP (Supplementary Fig. 9a). Furthermore, MC903 treatment resulted in the accumulation of basophils in the skin of wild-type mice (Supplementary Fig. 9b). MC903-induced basophil responses were TSLP-dependent, as treatment of wild-type mice with anti-TSLP diminished the accumulation of basophil populations (Supplementary Fig. 9b). MC903 treatment of *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ mice also resulted in a robust accumulation of basophils in the skin (Supplementary Fig. 9c), and MC903-induced, IL-3-independent basophil responses were significantly reduced after anti-TSLP treatment (Supplementary Fig. 9c). Collectively, these data demonstrate IL-3-independent, TSLP-dependent basophil responses *in vivo*. To test whether TSLP-dependent basophils contribute to the induction of $T_H2$



©2011 Macmillan Publishers Limited. All rights reserved

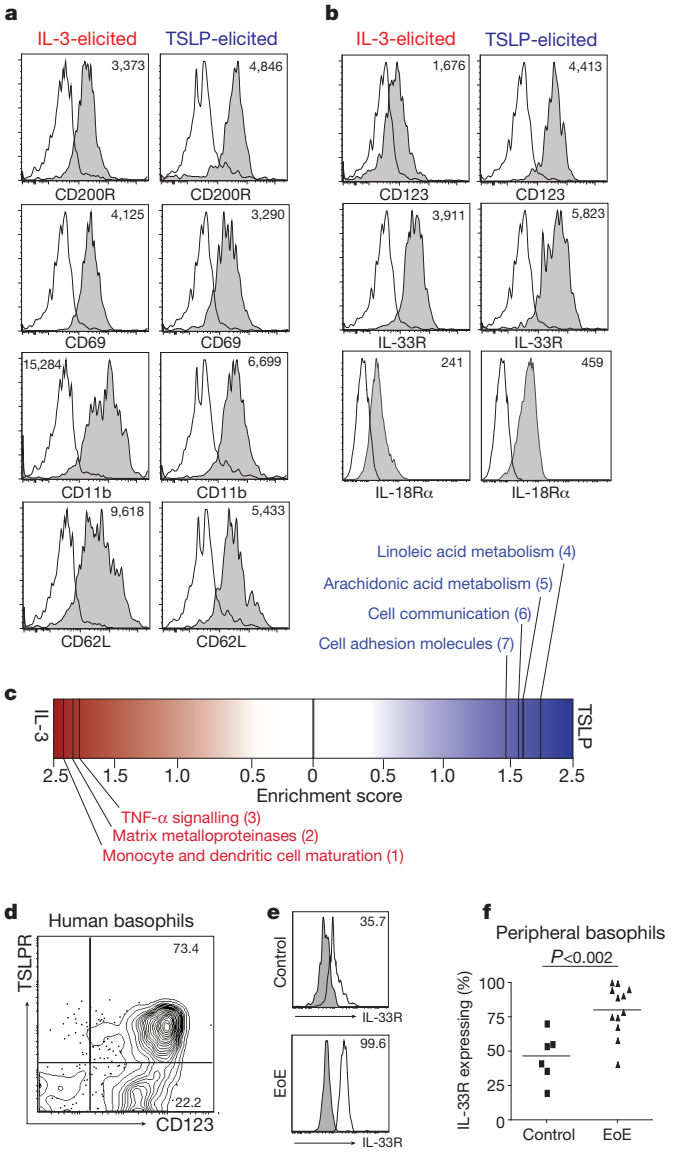



**Figure 4 | Murine and human basophil populations exhibit heterogeneity.** **a**, **b**, IL-3-elicited or TSLP-elicited murine basophils were stained (shaded) for CD200R, CD69, CD11b and CD62L (**a**), and CD123, IL-33R (T1/ST2) and IL-18Rα (**b**) expression and compared to fluorescence minus one controls. Numbers represent MFI levels. Results are representative of at least 3 separate experiments. **c**, GSEA of microarray data comparing IL-3-elicited and TSLP-elicited basophils. **d**, Activated human basophils stained for TSLPR. **e**, Representative histograms demonstrating IL-33R (ST2) expression (open) compared to isotype controls. **f**, Percentage of basophils from control or eosinophilic oesophagitis (EoE) patients expressing IL-33R compared to isotype controls. Statistical analysis was performed using a two-tailed Student's *t*-test.

cell responses, MC903 treatment of wild-type mice was extended to 7 days. MC903 treatment resulted in the robust accumulation of basophils in the skin (Supplementary Fig. 9d) and a marked increase in IL-4 and IL-5 production in the draining lymph nodes (Supplementary Fig. 9f). Critically, diphtheria-toxin-mediated depletion of basophils[30] reduced the accumulation of basophils in the skin (Supplementary Fig. 9e) and decreased IL-4 and IL-5 production (Supplementary Fig. 9f), demonstrating that TSLP-dependent basophils contribute to $T_H2$ cytokine responses *in vivo*.

To test whether human basophils can respond to TSLP, peripheral blood from healthy human donors was taken and TSLPR expression on basophils was evaluated. Greater than 70% of activated peripheral human basophils expressed TSLPR (Fig. 4d), indicating that human basophils can be TSLP-responsive. The food-allergy-associated disease eosinophilic oesophagitis is associated with gain-of-function polymorphisms in the gene encoding TSLP and elevated expression of TSLP[3–6,13]. To examine whether basophil heterogeneity exists when human basophils develop in a TSLP-rich environment associated with food allergy, basophil responses were examined in healthy control or eosinophilic oesophagitis patients. No differences in expression of HLA-DR, CD28, CD40, CD86, CD69 or CD203c were observed between control and eosinophilic oesophagitis patients (data not shown). However, consistent with TSLP-induced basophils in mice, basophils from eosinophilic oesophagitis patients expressed significantly higher levels of IL-33R compared to basophils from healthy controls (Fig. 4e, f). These data indicate that heterogeneity in basophil populations may exist in humans and correlates with susceptibility to allergic inflammation.

Collectively, these data identify that TSLP can selectively promote basophil haematopoiesis from bone-marrow-resident precursors and elicit mature basophil responses in the periphery in an IL-3–IL-3R-sufficient or -deficient environment. The demonstration that TSLP-elicited basophils exhibit divergent transcriptional and functional profiles from IL-3-elicited basophils indicates that similar to other immune cell lineages, basophils are a heterogeneous cell population that exhibit distinct phenotypic and functional characteristics depending on the cytokine environment in which they develop or are activated. The identification of a role for TSLP in promoting basophil haematopoiesis, peripheral basophilia and altering the functional capacity of basophils (Supplementary Fig. 10) may provide a biological mechanism through which local production of TSLP at one barrier surface can confer susceptibility to pan-allergic diseases at multiple mucosal sites.

## METHODS SUMMARY

C57BL/6 or BALB/c wild-type, IL-4/eGFP reporter or *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ mice were treated by intraperitoneal (i.p.) injection with PBS or 10 μg of rTSLP daily for 4–7 days and basophil populations were evaluated. Mice were injected intravenously (i.v.) with 10 μg of control or TSLP encoding cDNA plasmid and basophils (day 21) and small intestine pathology (day 28) were evaluated after injection. Wild-type and *Tslpr*$^{-/-}$ mice were infected with *T. muris* and on day 10 after infection, one group of *Tslpr*$^{-/-}$ mice received $100 \times 10^3$ TSLPR-sufficient basophils elicited from TSLP-cDNA-injected H2K-Bcl2-transgenic mice. All groups received 10 μg injections of rTSLP every other day starting at day 11 after infection. On day 21 after infection, $T_H2$ cytokine responses, serum IgE, pathology, RELM-β and worm burdens were evaluated. Bone marrow from wild-type, IL-4/eGFP reporter or *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ mice was cultured in the presence of 10 ng ml$^{-1}$ of IL-3 or 1 μg ml$^{-1}$ of TSLP for 5 days. Bone-marrow-derived basophils were sorted on day 5 after culture, RNA was isolated and gene expression profiles were determined using an Affymetrix HT Mouse 430 PM platform, and analysis was performed using GSEA software through http://www.broadinstitute.org/gsea/index.jsp. Bone-marrow-resident BaPs were sorted from wild-type whole bone marrow and cultured in the presence of 10 ng ml$^{-1}$ of IL-3 or 1 μg ml$^{-1}$ of TSLP for 5 days and basophils were evaluated by flow cytometry and cytology. Wild-type, *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$, or BaS-TRECK mice were treated with 50 μl of 100% ethanol or 50 μl of 2 nm MC903 dissolved in 100% ethanol topically for 3–7 days. Groups were treated (i.p.) with 1 mg of rat IgG isotype control or anti-TSLP antibody on day 0 of ethanol or anti-MC903 treatment. Skin-draining lymph node cells were isolated from ethanol-treated or MC903-treated mice on day 7 and stimulated with anti-CD3 and anti-CD28 for 48 h.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 24 May; accepted 27 June 2011.**
**Published online 14 August 2011.**

1. Anthony, R. M., Rutitzky, L. I., Urban, J. F. Jr, Stadecker, M. J. & Gause, W. C. Protective immune mechanisms in helminth infection. *Nature Rev. Immunol.* **7,** 975–987 (2007).
2. Umetsu, D. T., McIntire, J. J., Akbari, O., Macaubas, C. & DeKruyff, R. H. Asthma: an epidemic of dysregulated immunity. *Nature Immunol.* **3,** 715–720 (2002).
3. Harada, M. *et al. TSLP* promoter polymorphisms are associated with susceptibility to bronchial asthma. *Am. J. Respir. Cell Mol. Biol.* **44,** 787–793 (2011).



©2011 Macmillan Publishers Limited. All rights reserved

4. Sherrill, J. D. *et al.* Variants of thymic stromal lymphopoietin and its receptor associate with eosinophilic esophagitis. *J. Allergy Clin. Immunol.* **126,** 160–5e3 (2010).
5. Rothenberg, M. E. *et al.* Common variants at 5q22 associate with pediatric eosinophilic esophagitis. *Nature Genet.* **42,** 289–291 (2010).
6. Hunninghake, G. M. *et al.* TSLP polymorphisms are associated with asthma in a sex-specific fashion. *Allergy* **65,** 1566–1575 (2010).
7. Ying, S. *et al.* Thymic stromal lymphopoietin expression is increased in asthmatic airways and correlates with expression of Th2-attracting chemokines and disease severity. *J. Immunol.* **174,** 8183–8190 (2005).
8. Al-Shami, A. *et al.* A role for thymic stromal lymphopoietin in $CD4^+$ T cell development. *J. Exp. Med.* **200,** 159–168 (2004).
9. Ito, T. *et al.* TSLP-activated dendritic cells induce an inflammatory T helper type 2 cell response through OX40 ligand. *J. Exp. Med.* **202,** 1213–1223 (2005).
10. Taylor, B. C. *et al.* TSLP regulates intestinal immunity and inflammation in mouse models of helminth infection and colitis. *J. Exp. Med.* **206,** 655–667 (2009).
11. Mou, Z. *et al.* Overexpression of thymic stromal lymphopoietin in allergic rhinitis. *Acta Otolaryngol. (Stockh.)* **129,** 297–301 (2009).
12. Soumelis, V. *et al.* Human epithelial cells trigger dendritic cell mediated allergic inflammation by producing TSLP. *Nature Immunol.* **3,** 673–680 (2002).
13. Ziegler, S. F. & Artis, D. Sensing the outside world: TSLP regulates barrier immunity. *Nature Immunol.* **11,** 289–293 (2010).
14. Sebestyén, M. G. *et al.* Mechanism of plasmid delivery by hydrodynamic tail vein injection. I. Hepatocyte uptake of various molecules. *J. Gene Med.* **8,** 852–873 (2006).
15. Perrigoue, J. G. *et al.* MHC class II-dependent basophil-$CD4^+$ T cell interactions promote $T_H2$ cytokine-dependent immunity. *Nature Immunol.* **10,** 697–705 (2009).
16. Sokol, C. L. *et al.* Basophils function as antigen-presenting cells for an allergen-induced T helper type 2 response. *Nature Immunol.* **10,** 713–720 (2009).
17. Yoshimoto, T. *et al.* Basophils contribute to $T_H2$-IgE responses *in vivo* via IL-4 production and presentation of peptide-MHC class II complexes to $CD4^+$ T cells. *Nature Immunol.* **10,** 706–712 (2009).
18. Hammad, H. *et al.* Inflammatory dendritic cells—not basophils—are necessary and sufficient for induction of Th2 immunity to inhaled house dust mite allergen. *J. Exp. Med.* **207,** 2097–2111 (2010).
19. Ohnmacht, C. *et al.* Basophils orchestrate chronic allergic dermatitis and protective immunity against helminths. *Immunity* **33,** 364–374 (2010).
20. Tang, H. *et al.* The T helper type 2 response to cysteine proteases requires dendritic cell-basophil cooperation via ROS-mediated signaling. *Nature Immunol.* **11,** 608–617 (2010).
21. Zhou, B. *et al.* Thymic stromal lymphopoietin as a key initiator of allergic airway inflammation in mice. *Nature Immunol.* **6,** 1047–1053 (2005).
22. Gessner, A., Mohrs, K. & Mohrs, M. Mast cells, basophils, and eosinophils acquire constitutive IL-4 and IL-13 transcripts during lineage differentiation that are sufficient for rapid cytokine production. *J. Immunol.* **174,** 1063–1072 (2005).
23. Iwasaki, H. & Akashi, K. Myeloid lineage commitment from the hematopoietic stem cell. *Immunity* **26,** 726–740 (2007).
24. Lantz, C. S. *et al.* Role for interleukin-3 in mast-cell and basophil development and in immunity to parasites. *Nature* **392,** 90–93 (1998).
25. Scott, C. L. *et al.* Reassessment of interactions between hematopoietic receptors using common β-chain and interleukin-3-specific receptor β-chain-null cells: no evidence of functional interactions with receptors for erythropoietin, granulocyte colony-stimulating factor, or stem cell factor. *Blood* **96,** 1588–1590 (2000).
26. Haining, W. N. & Wherry, E. J. Integrating genomic signatures for immunologic discovery. *Immunity* **32,** 152–161 (2010).
27. Lenox, L. E. *et al.* Mutation of tyrosine 145 of lymphocyte cytosolic protein 2 protects mice from anaphylaxis and arthritis. *J. Allergy Clin. Immunol.* **124,** 1088–1098 (2009).
28. Siracusa, M. C., Perrigoue, J. G., Comeau, M. R. & Artis, D. New paradigms in basophil development, regulation and function. *Immunol. Cell Biol.* **88,** 275–284 (2010).
29. Li, M. *et al.* Induction of thymic stromal lymphopoietin expression in keratinocytes is necessary for generating an atopic dermatitis upon application of the active vitamin D3 analogue MC903 on mouse skin. *J. Invest. Dermatol.* **129,** 498–502 (2009).
30. Sawaguchi, M. *et al.* Basophils and mast cells differently regulate allergic response; *in vivo* demonstration by diphtheria toxin based deletion system. *Int. Congress Immunol.* WS/PP-043–04 (2011).


**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank members of the Artis laboratory for critical reading of the manuscript, A. Budelsky, T. Martin, B.-R. Park Yoon, J. Bigler and M. Timour (Amgen) for assistance with experiments and microarrays and the University of Pennsylvania flow cytometry core for assistance with sorting. Work in the Artis laboratory is supported by the National Institutes of Health (AI61570, AI74878, AI87990 and AI083480 (to D.A.), F32 fellowship AI085828 (to M.C.S.), F31 training grant GM082187 (to S.A.S), T32 training grant AI060516 (to D.A.H.)) and the Burroughs Wellcome Fund (to D.A.).

**Author Contributions** M.C.S., S.A.S., D.A.H., B.S.K., M.B.H., H.K.J., L.A.S., M.R.C. and D.A. designed and performed the research. S.F.Z., T.K. and M.K. provided reagents. E.C.D., A.C. and J.M.S. collected and provided human samples. M.C.S., S.A.S., D.A.H., B.S.K., T.A.D., E.J.W and D.A. analysed the data. M.C.S. and D.A. wrote the paper.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.A. (dartis@mail.med.upenn.edu).



©2011 **Macmillan Publishers Limited. All rights reserved**

## METHODS

**Mice.** BALB/c, C57BL/6 and *Tslpr*$^{-/-}$ mice were obtained from Charles River Laboratories. IL4/eGFP reporter mice, *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ mice, *Tslpr*$^{-/-}$ × *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ and H2K-Bcl2-transgenic mice and BaS-TRECK mice were bred at the University of Pennsylvania. The above strains were housed in specific pathogen-free conditions at the University of Pennsylvania. Control and SPC-TSLP mice were maintained at Benaroya Research Institute. All experiments were performed under Institutional Animal Care and Use Committee (IACUC) approved protocols and in accordance with the guidelines of the IACUC of the University of Pennsylvania. All mice used were 4–12 weeks of age. Wild-type, IL-4/eGFP reporter mice or *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ mice were treated by intraperitoneal injection with PBS or 10 μg of rTSLP daily for 4–7 days, and basophil populations were evaluated. Mice were injected (i.v.) with 10 μg of control or TSLP encoding cDNA plasmids (provided by M. Comeau), and basophil populations and small intestine pathology were evaluated on days 21 and 28 after injection. *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ mice were given intraperitoneal injections of anti-TSLP (1 mg) on day 0 and 4 and basophil populations were evaluated on day 7 as previously described[10,15,31,32].

**Recombinant cytokines.** rmTSLP (R&D Systems, M. Comeau, Amgen), rmIL-3 (R&D Systems, M. Comeau, Amgen), rmSCF (R&D Systems), rmIL-18 (R&D Systems), rhIL-3 (M. Comeau, Amgen).

**Flow cytometry.** Murine spleens were homogenized by passing through a 70-μm nylon mesh filter, and stained with anti-mouse fluorochrome-conjugated monoclonal antibodies against CD3ε, CD4, CD8, B220, CD19, CD11b, CD49b, FcεRIα, c-Kit, CD123, CD69, CD11b, IL-18Rα and CD62L (eBioscience or BD Bioscience). Fluorochrome-conjugated monoclonal antibodies against CD200R and IL-33R were obtained from AbD serotec and mdbioproducts. Bone-marrow-resident cells were collected by removing and sterilizing femurs in 70% ethanol, removing the ends of the femurs and flushing cells out with 10 ml of complete media. Human peripheral blood cells were stained with anti-human fluorochrome-conjugated monoclonal antibodies against CD203c, HLA-DR, CD28, CD40, CD86, CD69, TSLPR, IgE, FcεRI, CD19 and TCRα/β (eBioscience, Miltenyi Biotec and mdbioproducts). Cells were run on a BD LSR II flow using DiVa software (BD Bioscience) and analysed with FlowJo software (version 8.7.1; Tree Star).

***T. muris* infection and TSLPR-sufficient basophil transfer.** Wild-type and *Tslpr*$^{-/-}$ mice were infected with 200 *Trichuris muris* eggs via oral gavage as previously described[10,15,31,32], and on day 21 after infection basophil populations were quantified. Wild-type and *Tslpr*$^{-/-}$ mice were infected with 200 *T. muris* eggs via oral gavage and on day 10 after infection one group of *Tslpr*$^{-/-}$ mice received $100 \times 10^3$ TSLPR-sufficient TSLP-cDNA-elicited basophils from H2K-Bcl2-transgenic mice. H2K-Bcl2-transgenic basophils were used to assure cell survival after adoptive transfer. All groups received 10 μg injections of rTSLP every other day starting at day 11 after infection. On day 21 after infection, *T. muris* antigen-specific $T_H2$ cytokine responses, serum IgE, pathology and worm burdens were evaluated as previously described[10,15,31,32].

**Bone-marrow-derived basophils.** Bone-marrow-resident cells from wild-type, IL-4/eGFP reporter mice or *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ mice were cultured in the presence of 10 ng ml$^{-1}$ of IL-3 or 1 μg ml$^{-1}$ of TSLP for 5 days, and basophil populations were evaluated by flow cytometry using a BD LSR II flow cytometer or sort-purified using a BD FACS Aria II flow cytometer. RNA from sort-purified bone-marrow-derived basophils was isolated, cDNA was made and gene expression profiles were determined by real time PCR (Qiagen, SYBR Green) or genome-wide transcriptional profiling using an Affymetrix HT Mouse 430 PM platform. Microarray analysis was performed using GSEA software (http://www.broadinstitute.org/gsea/index.jsp).

**Bone-marrow-derived basophil stimulation assays.** Bone-marrow-derived sort-purified basophils were stimulated with media, 10 ng ml$^{-1}$ of IL-3, 20 ng ml$^{-1}$ of IL-18 or 10 ng ml$^{-1}$ of IL-33, and cell-free supernatants were analysed using a multi-analyte panel (Rodent v2.0) (http://www.rulesbasedmedicine.com/). Bone-marrow-derived basophils were grown in 10 ng ml$^{-1}$ of IL-3 or 1 μg ml$^{-1}$ of TSLP for 5 days and sort-purified. Bone-marrow-derived mast cells were grown in IL-3 (10 ng ml$^{-1}$) and SCF (12.5 ng ml$^{-1}$) for 8 weeks. IL-3-elicited basophils and mast cells were allowed to rest overnight in media containing 10 ng ml$^{-1}$ of IL-4 and 10 ng ml$^{-1}$ of IL-3. TSLP-elicited basophils were rested in media containing 10 ng ml$^{-1}$ of IL-4 and 1 μg ml$^{-1}$ of TSLP. Cells were incubated with 1 μg ml$^{-1}$ anti-dinitrophenol IgE for 1 h at 37 °C, washed, and stimulated with 100 ng ml$^{-1}$ dinitrophenol-conjugated human serum albumin (HSA-DNP; Sigma) for 1 h. Supernatants were tested for β-hexosaminidase activity by incubating with 1 mM *p*-nitrophenyl-*N*-acetyl-β-D-glucosamide. Absorbance was read at 405 nm. Serial dilutions of a known quantity of lysed bone-marrow-derived mast cells were used as a standard reference.

**Basophil survival assays.** Basophils were sort-purified from the blood and spleens of wild-type mice and cultured in the presence 10 ng ml$^{-1}$ of IL-3 or 1 μg ml$^{-1}$ of TSLP for 24 h, and cell viability was determined by 7-AAD (BD Bioscience) exclusion.

**Basophil progenitor cultures.** Basophil progenitor cells (identified as NBNT, CD34$^+$, FceRI$^+$, c-Kit$^-$) were sort-purified from whole bone marrow isolated from wild-type mice and cultured in the presence of 10 ng ml$^{-1}$ of IL-3 or 1 μg ml$^{-1}$ of TSLP for 5 days, and basophil populations were evaluated by flow cytometry and cytology.

**MC903-induced atopic dermatitis.** Wild-type or *Csf2rb2*$^{-/-}$*Csf2rb*$^{-/-}$ mice were treated with 50 μl of 100% ethanol or 50 μl of 2 nM MC903 dissolved in 100% ethanol topically for 3–7 days. Groups were treated (i.p.) with 1 mg of rat IgG isotype control or anti-TSLP neutralizing antibody on day 0. BaS-TRECK mice were treated for 7 days with MC903 and given intraperitoneal injections of diphtheria toxin (750 ng) on day 0 and day 4 after treatment. Skin was mechanically disrupted, single-cell suspensions were prepared and basophil populations were evaluated by flow cytometry. Ears were separated into two halves and floated dermis side down in complete media for 24 h to determine local production of TSLP, and serum was collected to determine systemic TSLP levels by ELISA (eBioscience). Skin-draining lymph node cells were isolated from ethanol-treated or MC903-treated mice on day 7 and stimulated with anti-CD3 and anti-CD28 for 48 h, and IL-4 and IL-5 production was determined by ELISA (eBioscience).

**Human patients and sample preparation.** All subjects were studied under IRB approval from The Children's Hospital of Philadelphia. All subjects signed an informed consent. EoE was defined by the current consensus guidelines of having greater than 15 eosinophils per high powered field (HPF) at ×400 light microscopy isolated to the oesophagus after other diagnoses were excluded and after being at least 8 weeks on optimal dose of proton pump inhibitors. Human blood was collected from healthy donors and peripheral blood cells were isolated using a Ficoll gradient. Basophils were isolated using a MACS basophil isolation kit II according to the manufacturer's protocol (Miltenyi Biotec, 130-092-662). Basophils were cultured overnight in 50 ng ml$^{-1}$ of IL-3 and stained for TSLPR. Human blood was collected from healthy paediatric donors and paediatric donors diagnosed with eosinophilic oesophagitis. Peripheral blood cells were isolated by Ficoll gradient, and surface marker expression was determined by flow cytometry.

**Statistics.** Results are shown as means ± s.e.m. for individual animals. Statistical significance was determined by Student's *t*-test. Results were considered significant at $P < 0.05$.


31. Zaph, C. *et al.* Epithelial-cell-intrinsic IKK-β expression regulates intestinal immune homeostasis. *Nature* **446,** 552–556 (2007).
32. Nair, M. G. *et al.* Goblet cell-derived resistin-like molecule β augments CD4+ T cell production of IFN-γ and infection-induced intestinal inflammation. *J. Immunol.* **181,** 4709–4715 (2008).


©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10332

# $N_2O$ binding at a [4Cu:2S] copper–sulphur cluster in nitrous oxide reductase

Anja Pomowski[1], Walter G. Zumft[2], Peter M. H. Kroneck[3] & Oliver Einsle[1,4]

**Nitrous oxide ($N_2O$) is generated by natural and anthropogenic processes and has a critical role in environmental chemistry. It has an ozone-depleting potential similar to that of hydrochlorofluorocarbons as well as a global warming potential exceeding that of $CO_2$ 300-fold[1,2]. In bacterial denitrification, $N_2O$ is reduced to $N_2$ by the copper-dependent nitrous oxide reductase ($N_2OR$)[3]. This enzyme carries the mixed-valent $Cu_A$ centre and the unique, tetranuclear $Cu_Z$ site. Previous structural data were obtained with enzyme isolated in the presence of air that is catalytically inactive without prior reduction. Its $Cu_Z$ site was described as a [4Cu:S] centre, and the substrate-binding mode and reduction mechanism remained elusive. Here we report the structure of purple $N_2OR$ from *Pseudomonas stutzeri*, handled under the exclusion of dioxygen, and locate the substrate in $N_2O$-pressurized crystals. The active $Cu_Z$ cluster contains two sulphur atoms, yielding a [4Cu:2S] stoichiometry; and $N_2O$ bound side-on at $Cu_Z$, in close proximity to $Cu_A$. With the substrate located between the two clusters, electrons are transferred directly from $Cu_A$ to $N_2O$, which is activated by side-on binding in a specific binding pocket on the face of the [4Cu:2S] centre. These results reconcile a multitude of available biochemical data on $N_2OR$ that could not be explained by earlier structures, and outline a mechanistic pathway in which both metal centres and the intervening protein act in concert to achieve catalysis. This structure represents the first direct observation, to our knowledge, of $N_2O$ bound to its reductase, and sheds light on the functionality of metalloenzymes that activate inert small-molecule substrates. The principle of using distinct clusters for substrate activation and for reduction may be relevant for similar systems, in particular nitrogen-fixing nitrogenase[4].**

$N_2O$ is a kinetically inert, colourless gas that is being scrutinized for its deleterious effects on global climate and, in this context, was named the most critical anthropogenic substance emitted in the twenty-first century[1]. Although the net amount of atmospheric $CO_2$ is significantly higher, its anthropogenic emissions have been put under regulation and control, whereas $N_2O$ concentrations are on a steady rise[1,5]. $N_2O$ is generated by abiotic processes, for instance in hypersaline ponds in the Antarctic[6] that are subject to major fluctuations with an impact on global climate[7]. It also occurs as a significant by-product of the bacterial respiratory pathways of nitrification and denitrification that are strongly enhanced by fertilization in modern agriculture[8,9], leading to an overall imbalance in the global nitrogen cycle[5]. Although the reduction of $N_2O$ to dinitrogen—$N_2$—is strongly exergonic ($N_2O + 2\,H^+ + 2\,e^- \rightarrow N_2 + H_2O$; $\Delta G^{0\prime} = -339.5\,kJ\,mol^{-1}$), electronic delocalization stabilizes the molecule and leads to an activation energy barrier of $250\,kJ\,mol^{-1}$ (ref. 10). Its reduction by bacteria is catalysed by $N_2OR$ (ref. 3), a copper-dependent enzyme that forms tightly linked head-to-tail homodimers of 130 kDa (Fig. 1). $N_2OR$ contains two multi-copper centres, $Cu_A$ and $Cu_Z$, with distinct spectroscopic properties. $Cu_A$ is a mixed-valent [$Cu_{A1}^{+1.5}$:$Cu_{A2}^{+1.5}$] S = 1/2 centre that is also found in some haem-copper oxidases[11] and was engineered into the type I Cu site of azurin[12]. It was suggested to transfer one electron from an external donor such as cytochrome *c*-552 to the active site at $Cu_Z$[13]. The three-dimensional structure of $N_2OR$ from *Marinobacter hydrocarbonoclasticus* (formerly *Pseudomonas nautica*)[14] revealed $Cu_Z$ to be tetranuclear, and it was subsequently recognized to be a $\mu_4$-sulphido-bridged [4Cu:S] centre[15]. This was confirmed by structures of $N_2OR$ from *Paracoccus denitrificans*[16] and '*Achromobacter cycloclastes*'[17], but no direct evidence was obtained concerning the mode of substrate binding or the mechanism of its conversion. One[16] or two[17] water ligands were observed bound to a face of the $Cu_Z$ cluster between atoms $Cu_{Z1}$ and $Cu_{Z4}$ and theoretical considerations have suggested the binding of the substrate $N_2O$ in a bridging manner in this position[18,19].

$N_2OR$ was identified in 1982[20], and a wealth of biochemical data has been accumulated since. However, seeming inconsistencies arose

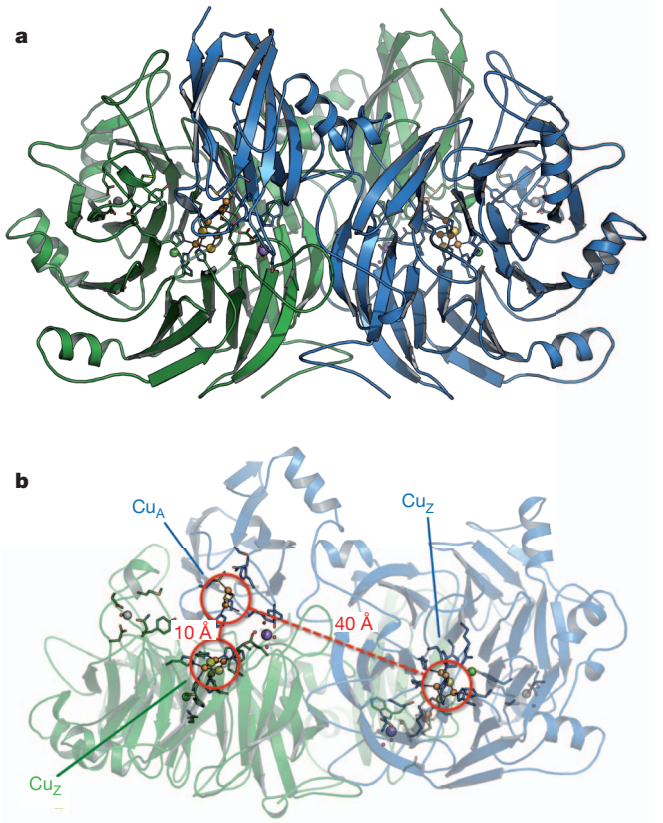



**Figure 1 | $N_2OR$ from *Pseudomonas stutzeri*.** **a**, The enzyme is a homodimer in head-to-tail orientation with the tetranuclear $Cu_Z$ active site located in the N-terminal, seven-bladed β-propeller domain and the binuclear $Cu_A$ site in the C-terminal cupredoxin domain. **b**, The $Cu_A$–$Cu_Z$ distance within one monomer is around 40 Å and thus too large for efficient electron transfer. Instead, the $Cu_A$ site of one monomer is at a distance of only 10 Å from $Cu_Z$ in the other monomer.

[1]Lehrstuhl für Biochemie, Institut für organische Chemie und Biochemie, Albert-Ludwigs-Universität Freiburg, Albertstr. 21, 79104 Freiburg, Germany. [2]Institut für Angewandte Biowissenschaften, Abteilung Molekulare Mikrobiologie, Karlsruher Institut für Technologie, Hertzstr. 16, 76187 Karlsruhe, Germany. [3]Fachbereich Biologie, Universität Konstanz, Universitätsstr. 10, 78457 Konstanz, Germany. [4]Centre for Biological Signalling Studies (bioss), Albert-Ludwigs-Universität Freiburg, Hebelstr. 25, 79104 Freiburg, Germany.



©2011 Macmillan Publishers Limited. All rights reserved

from the fact that studies were carried out on two different forms of the enzyme[3,21,22]. $N_2OR$ prepared under anoxic conditions differs from enzyme isolated in the presence of dioxygen, in particular with respect to the $Cu_Z$ site. An active, purple form of the enzyme (form I) was obtained in anoxic preparations from *P. stutzeri*[21,22] and from *A. cycloclastes*[23]. Oxic preparations produce a pink enzyme (form II) with only 20–40% of the specific activity of form I (ref. 3). However, when fully reduced to the all-Cu(I) state[19], form II $N_2OR$ shows a specific activity for $N_2O$ reduction of 4.5 μkat $mg^{-1}$ that is fourfold higher than the one of *P. stutzeri* $N_2OR$ form I (1.1 μkat $mg^{-1}$)[21]. Reductive activation of $N_2OR$ II requires methyl viologen, at a redox potential that is not achieved in the bacterial periplasm, while the all-Cu(I) state of *P. stutzeri* $N_2OR$ I was not obtained even after 65 h of dithionite incubation[22]. We therefore argue that the modes of substrate activation and electron transfer of $N_2OR$ I are distinct from what was suggested for all-Cu(I) $N_2OR$ II[17,19,24].

Although overall the structure of anoxically isolated *P. stutzeri* $N_2OR$ I is very similar to the structures reported thus far, both metal centres show marked differences. $Cu_A$ is of a characteristic, pink colour in the mixed-valent $[Cu_{A1}{}^{+1.5}{:}Cu_{A2}{}^{+1.5}]$ $S = 1/2$ state and its conformation was virtually identical in its oxidized form in *A. cycloclastes* $N_2OR$[17] and in the reduced form in *M. hydrocarbonoclasticus*[14] and *P. denitrificans*[16] $N_2OR$. The spectroscopic hallmark of $Cu_A$ is the narrow 7-line hyperfine splitting in the $g_{||}$ region of the electron paramagnetic resonance (EPR) spectrum (Supplementary Fig. 6) that originates from an unpaired electron fully delocalized over two copper nuclei with a nuclear spin $I_{Cu} = 3/2$ (refs 11, 25). Each copper has a single histidine ligand, His 583 for $Cu_{A1}$ and His 626 for $Cu_{A2}$ in *P. stutzeri* (Fig. 2b). In previous descriptions of the centre, the histidine ligand to $Cu_{A1}$ coordinated the metal through its $N_{\delta1}$ atom at a canonical Cu–$N_{His}$ distance of approximately 2.1 Å. However, in purple $N_2OR$ from *P. stutzeri*, H583 was not a ligand to $Cu_{A1}$. Instead, the imidazole ring of H583 was rotated by approximately 130° to form a short (2.6 Å) hydrogen bond between the $N_{\delta1}$ atom of H583 and the $O_\gamma$ atom of S550 and a second one at 2.8 Å from the $N_{\varepsilon2}$ atom of the same histidine to the β-carboxy group of D576 (Fig. 2b and Supplementary Fig. 1a). As a consequence, the now three-coordinate $Cu_{A1}$ atom was retracted into the plane formed by the sulphur atoms of C618, C622 and M629. The $Cu_{A1}$–$Cu_{A2}$ distance was 2.5 Å, and this protein still showed the characteristic, 7-line EPR signal (Supplementary Fig. 6). Both S550 and D576 are conserved residues among $N_2O$ reductases with D576 forming a putative entry point for electrons. This arrangement suggests a mechanistic relevance, in that reduction of $Cu_A$ from an external electron donor is only made possible after H583 flips back to ligate $Cu_{A1}$, and this rearrangement occurred only after exposure to $N_2O$ (see later).

The second cluster—$Cu_Z$—is coordinated by seven histidine residues originating from six of the seven blades of a β-propeller domain (Fig. 1). In *P. stutzeri* $N_2OR$ I the ligands to $Cu_Z$ are H129, H130, H178, H326, H382, H433 and H494. The occurrence of positive- or negative-difference electron density was used in modelling, and with slightly reduced occupancies (see Supplementary Discussion) the model of the previously observed $\mu_4$-[4Cu:S] cluster was well accommodated. However, the maps revealed electron density for an additional ligand bridging $Cu_{Z1}$ and $Cu_{Z4}$. Although in earlier crystallographic studies similar species were modelled as water[16,17,26], a water molecule at this position in *P. stutzeri* $N_2OR$ resulted in a positive-difference electron density peak (Supplementary Fig. 7a). The atom refined to low *B* factors, indicating that a more electron-rich species was required. A second sulphur atom—$S_{Z2}$—optimally explained the observed electron density (Supplementary Fig. 3c). We thus propose a composition of [4Cu:2S] for the $Cu_Z$ cluster in $N_2OR$ I (Fig. 2a and Supplementary Fig. 1b). This interpretation is supported by several experimental

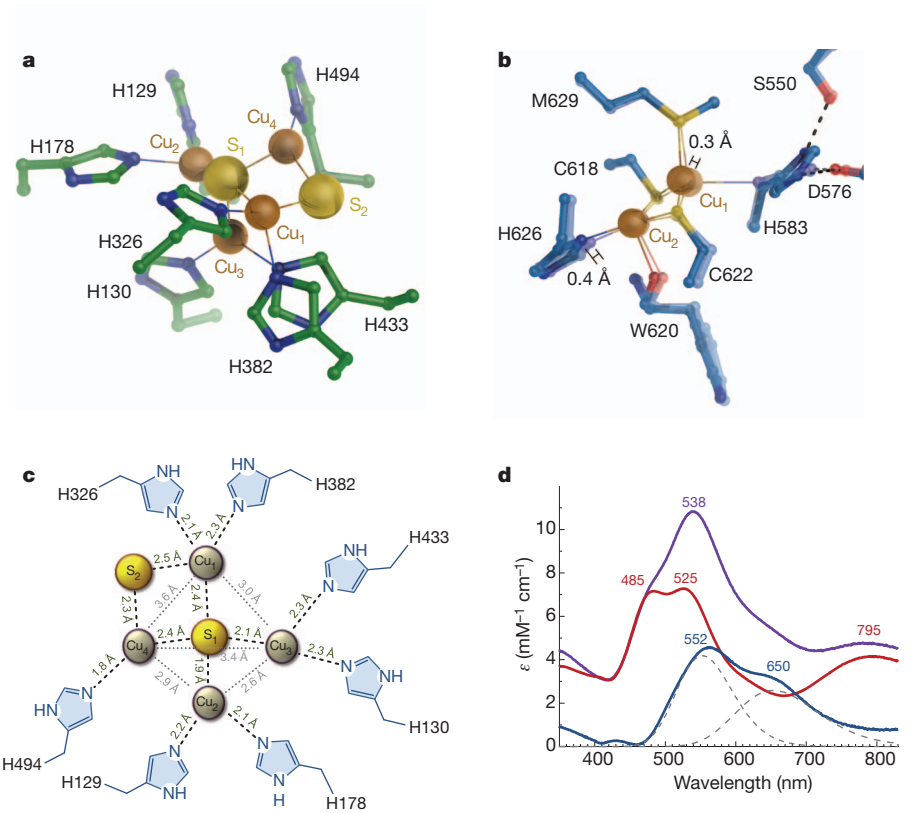



**Figure 2 | The copper sites of purple *P. stutzeri* $N_2OR$.** **a**, The $Cu_Z$ site containing two sulphur atoms ($S_{Z1}$ and $S_{Z2}$). **b**, The $Cu_A$ site, overlaying the state with His 583 ligating $Cu_{A1}$ (transparent) and the as isolated form (opaque), where the histidine is not a ligand to the metal. **c**, Bond distances in and around the $Cu_Z$ site. **d**, Ultraviolet/visible light spectrum of *P. stutzeri* $N_2OR$ (purple). The spectrum can be deconvoluted into the contributions of the $Cu_A$ site (red) and the tetranuclear $Cu_Z$ site (blue). The signature spectrum of $Cu_Z$ consists of two individual bands (grey) that represent distinct charge-transfer transitions from the two sulphur atoms. For stereo images of **a**, **b** see Supplementary Fig. 1.



©2011 Macmillan Publishers Limited. All rights reserved

observations. First, the observed bond distances to $Cu_{Z1}$ (2.6 Å) and $Cu_{Z4}$ (2.5 Å) are closer to typical Cu–S coordinative bonds than to Cu–O bonds (Fig. 2c). Second, in *A. cycloclastes* $N_2OR$ an iodide anion that was added as a specific inhibitor of enzymatic activity binds in the same position as $S_{Z2}$ does in *P. stutzeri* $N_2OR$ I[17] and $I^-$, as a strong ligand according to the spectrochemical series, is well suited to replace $S^{2-}$[27]. Third, the electron excitation spectrum of *P. stutzeri* $N_2OR$ can be deconvoluted into the individual contributions of the metal sites $Cu_A$ and $Cu_Z$ (Fig. 2d; see Supplementary Discussion and Supplementary Figs 3 and 4), and the $Cu_Z$ site of $N_2OR$ I can be modelled by two distinct maxima at 552 nm and 650 nm (Supplementary Fig. 3c). They fall within an energy range typical for S→Cu charge transfer transitions involving mainly Cu 3*d* and sulphur 3*p* orbitals[28] and we interpret them to originate from the two sulphur atoms of $Cu_Z$, $S_{Z1}$ and $S_{Z2}$, respectively.

The tetranuclear site of $N_2OR$ was described as a mixture of an active form ($Cu_Z$) and an inactive form ($Cu_Z$*)[3], and our structural data reflects this difference: whereas in $N_2OR$ I the active [4Cu:2S] $Cu_Z$ dominates, the inactive [4Cu:S] $Cu_Z$* is the main species in $N_2OR$ II, with the residual activity in the as isolated state being due to a small remaining fraction of $Cu_Z$. Atom $S_{Z2}$ is directly exposed to the substrate channel, and in oxic preparations, $O_2$ may diffuse to the $Cu_Z$ site and remove the bridging sulphur $S_{Z2}$, concomitant with the loss of enzymatic activity, as $Cu_Z$ irreversibly converts to the [4Cu:S] core cluster $Cu_Z$*. Correspondingly, the 552 nm band is lost in the spectral signature of $Cu_Z$*. Removing this contribution from a spectrum of the purple form I (Fig. 2d and Supplementary Fig. 4a) leads to a spectrum highly similar to that of *A. cycloclastes* $N_2OR$ II (Supplementary Fig. 4b)[17]. The origin of the 650 nm band is the $S_{Z1}$ atom. This band only disappears when all surrounding copper ions are reduced to the Cu(I) state. In consequence, the sensitivity of the $S_{Z2}$ atom is the Achilles' heel of the enzyme and the reason for its frequent loss of activity in microoxic environments.

The physiologically active state of the enzyme should be able to bind the substrate $N_2O$. Consequently, we pressurized crystals of $N_2OR$ I with $N_2O$ gas, followed by flash cooling in liquid nitrogen and X-ray diffraction data collection. In individual observations of $Cu_Z$, this resulted in the appearance of an elongated difference electron density feature (Fig. 3c) that was adequately modelled with a linear $N_2O$ molecule. At equal, fixed occupancies, the observed *B* factors of both the $N_2O$ molecule and the $Cu_Z$ centre were very close to those of the surrounding protein. Although the precise positioning of the substrate on the cluster showed slight variations, the overall orientation of all observed $N_2O$ molecules was consistent (see Supplementary Discussion). Notably, in all eight independent copies of $Cu_A$ in the asymmetric unit, ligand H583 was now rotated back to ligate $Cu_{A1}$ weakly, at a distance of 2.6 Å.

Little experimental evidence and only few theoretical calculations are available concerning the interaction of $N_2O$ with metal sites[10,29]. A mechanistic study based on the structure of a reduced form of $N_2OR$ isolated in the presence of dioxygen suggested that the $Cu_{Z1}$–$Cu_{Z4}$ edge of the cluster was the site of substrate binding[18], but the discovery of a second sulphur atom in the $Cu_Z$ centre of purple $N_2OR$ contrasts with this suggestion. Instead, we observe $N_2O$ positioned at $Cu_Z$ in a side-on manner on the cluster face built by the atoms $Cu_{Z2}$, $Cu_{Z4}$ and $S_{Z1}$ (Fig. 3). This is determined in part by the intricate arrangement of a hydrophobic substrate channel leading from the protein surface to the $Cu_Z$ cluster (Fig. 3a and Supplementary Fig. 2). The channel terminates close to $Cu_{Z1}$ and forms a proximal vestibule of sufficient size for the reorientation of the substrate $N_2O$. Note that F621 and M627, both part of the $Cu_A$ binding loop in the cupredoxin domain of the neighbouring monomer, form a barrier at this point. In the $N_2O$-bound structure the two residues move slightly apart as compared to the substrate-free enzyme, to accommodate the $N_2O$ molecule. We hypothesize that an initial interaction of the more polar end of $N_2O$ (the O atom) with the $S_\delta$ sulphur atom of M627 is the selection mechanism assuring the correct orientation of the substrate. Above the face of the $Cu_Z$ cluster $N_2O$ resides within a tight binding pocket formed by F621, H626 (a ligand to $Cu_A$) and M627 adjacent to a larger, distal cavity where several coordinated water molecules are located (Fig. 3a and Supplementary Fig. 2). The architecture of the substrate channel thus allows the apolar $N_2$, end product of the $N_2OR$ reaction, to leave the enzyme through the hydrophobic channel to the surface, while $H_2O$ is retained in the distal, water-filled cavity within the protein matrix. The observed distance between $N_2O$ and $Cu_Z$ suggests a weak

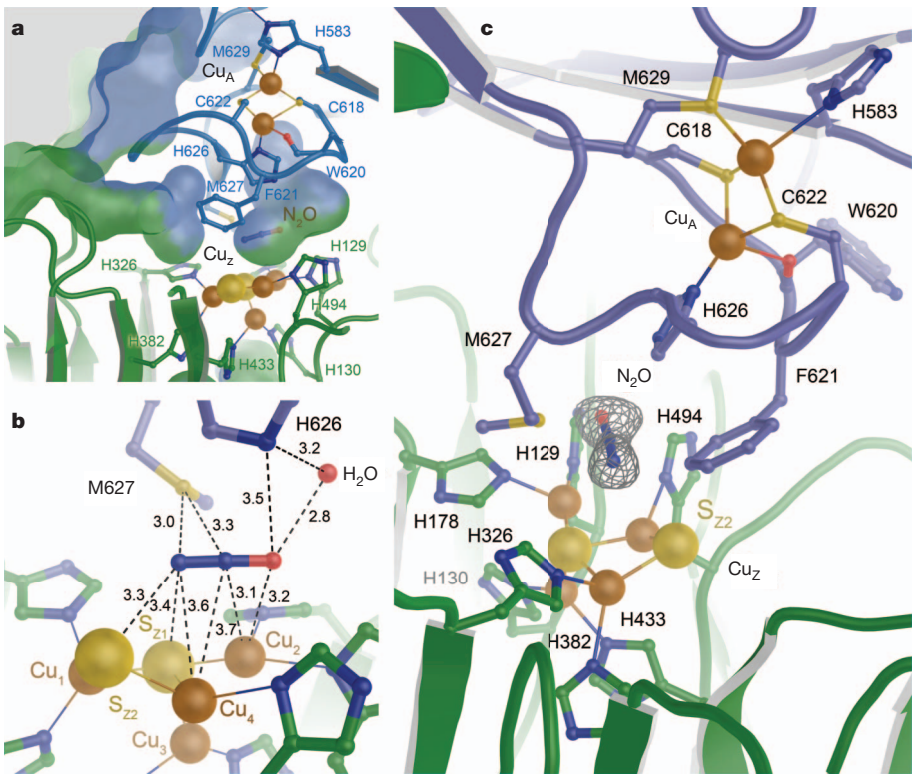



**Figure 3 | Substrate binding in $N_2OR$.** **a**, Active-site substrate binding. A hydrophobic substrate channel leads from the protein surface (left) to a proximal vestibule at $Cu_Z$, where the linear $N_2O$ molecule can re-orient to displace the two residues shielding the cluster, F621 and M627. The substrate enters with the oxygen atom first and is then bound in a side-on manner on the face of the cluster. The reduction takes place in a hydrophilic (water-filled), distal chamber where the reaction by-product $H_2O$ remains while the product $N_2$ exits via the hydrophobic entry channel. **b**, Distances around the $N_2O$ ligand (in Å). The closest interaction partners are H626, M627 and the $Cu_Z$ cluster. **c**, A $F_o - F_c$ difference electron density map contoured at the 3σ level showed the presence of the substrate. For stereo images of **a**, **b** see Supplementary Fig. 2.



©2011 Macmillan Publishers Limited. All rights reserved

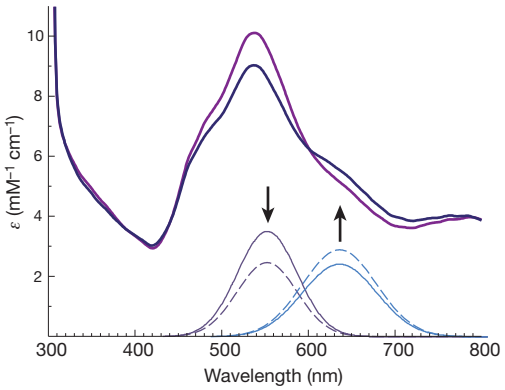



**Figure 4 | Changes in the electron excitation spectrum of $N_2$OR form I upon binding of the substrate $N_2$O.** The $N_2$OR I spectrum is shown in purple; the $N_2$O-bound form in blue. On the basis of the deconvolution of the spectra (Supplementary Fig. 3) the difference can be explained entirely by alterations of the two transitions that constitute the $Cu_Z$ spectrum (Fig. 2c and Supplementary Fig. 4a), a decrease of the 552 nm band concomitant with an increase of the 650 nm band. Arrows signify the increase/decrease of the bands upon transition from the purple to the blue curve when $N_2$O binds.

interaction, supported by EPR and ultraviolet/visible light spectra (Fig. 4 and Supplementary Figs 5 and 6). For the process of activation and reduction of $N_2$O, a stronger interaction might be required, which would involve a conformational change in the protein that is yet to be identified. The positioning of $N_2$O between the two metal sites points towards a reaction mechanism where the substrate, precisely oriented by the protein, is activated by the $Cu_Z$ cluster. Only after binding of $N_2$O is electron transfer from $Cu_A$ to $Cu_Z$ observed (Supplementary Fig. 5), and thus, rather than being a mere electron transfer site, $Cu_A$ of $N_2$OR is an integral part of the enzyme's catalytic centre.

## METHODS SUMMARY

*P. stutzeri* $N_2$OR was isolated and crystallized under strict exclusion of dioxygen as described previously[30]. Crystals of $N_2$OR were obtained by sitting-drop vapour diffusion. For the formation of substrate complexes, crystals were pressurized with 15 bar of $N_2$O gas for 15 min in a pressure cell and subsequently flash frozen in liquid nitrogen for diffraction data collection.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 17 September 2010; accepted 23 June 2011.**
**Published online 14 August 2011.**

1. Ravishankara, A. R., Daniel, J. S. & Portmann, R. W. Nitrous oxide ($N_2$O): the dominant ozone-depleting substance emitted in the 21st century. *Science* **326,** 123–125 (2009).
2. Crutzen, P. J., Mosier, A. R., Smith, K. A. & Winiwarter, W. $N_2$O release from agro-biofuel production negates global warming reduction by replacing fossil fuels. *Atmos. Chem. Phys.* **8,** 389–395 (2008).
3. Zumft, W. G. & Kroneck, P. M. H. Respiratory transformation of nitrous oxide ($N_2$O) to dinitrogen by Bacteria and Archaea. *Adv. Microb. Physiol.* **52,** 107–227 (2006).
4. Einsle, O. *et al.* Nitrogenase MoFe-protein at 1.16 Å resolution: a central ligand in the FeMo-cofactor. *Science* **297,** 1696–1700 (2002).
5. Canfield, D. E., Glazer, A. N. & Falkowski, P. G. The evolution and future of Earth's nitrogen cycle. *Science* **330,** 192–196 (2010).
6. Samarkin, V. A. *et al.* Abiotic nitrous oxide emission from the hypersaline Don Juan Pond in Antarctica. *Nature Geosci.* **3,** 341–344 (2010).
7. Schmittner, A. & Galbraith, E. D. Glacial greenhouse-gas fluctuations controlled by ocean circulation changes. *Nature* **456,** 373–376 (2008).
8. Einsle, O. & Kroneck, P. M. H. Structural basis of denitrification. *Biol. Chem.* **385,** 875–883 (2004).
9. Zumft, W. G. Cell biology and molecular basis of denitrification. *Microbiol. Mol. Biol. Rev.* **61,** 533–616 (1997).
10. Tolman, W. B. Binding and activation of $N_2$O at transition-metal centers: recent mechanistic insights. *Angew. Chem. Int. Ed.* **49,** 1018–1024 (2010).
11. Kroneck, P. M. H., Antholine, W. A., Riester, J. & Zumft, W. G. The nature of the cupric site in nitrous oxide reductase and of $Cu_A$ in cytochrome *c* oxidase. *FEBS Lett.* **248,** 212–213 (1989).
12. Savelieff, M. G. & Lu, Y. $Cu_A$ centers and their biosynthetic models in azurin. *J. Biol. Inorg. Chem.* **15,** 461–483 (2010).
13. Rasmussen, T., Brittain, T., Berks, B. C., Watmough, N. J. & Thomson, A. J. Formation of a cytochrome c–nitrous oxide reductase complex is obligatory for $N_2$O reduction by *Paracoccus pantotrophus*. *Dalton Transact.* **21,** 3501–3506 (2005).
14. Brown, K. *et al.* A novel type of catalytic copper cluster in nitrous oxide reductase. *Nature Struct. Biol.* **7,** 191–195 (2000).
15. Rasmussen, T. *et al.* The catalytic center in nitrous oxide reductase, $Cu_Z$, is a copper–sulfide cluster. *Biochemistry* **39,** 12753–12756 (2000).
16. Haltia, T. *et al.* The crystal stucture of nitrous oxide reductase ($N_2$OR) from *Paracoccus denitrificans* at 1.6 Å resolution. *Biochem. J.* **369,** 77–88 (2003).
17. Paraskevopoulos, K., Antonyuk, S. V., Sawers, R. G., Eady, R. R. & Hasnain, S. S. Insight into catalysis of nitrous oxide reductase from high-resolution structures of resting and inhibitor-bound enzyme from *Achromobacter cycloclastes*. *J. Mol. Biol.* **362,** 55–65 (2006).
18. Gorelsky, S. I., Ghosh, S. & Solomon, E. I. Mechanism of $N_2$O reduction by the $\mu_4$-S tetranuclear $Cu_Z$ cluster of nitrous oxide reductase. *J. Am. Chem. Soc.* **128,** 278–290 (2006).
19. Ghosh, S. *et al.* Activation of $N_2$O reduction by the fully reduced $\mu_4$-sulfide bridged tetranuclear $Cu_Z$ cluster in nitrous oxide reductase. *J. Am. Chem. Soc.* **125,** 15708–15709 (2003).
20. Zumft, W. G. & Matsubara, T. A novel kind of multi-copper protein as terminal oxidoreductase of nitrous-oxide respiration in *Pseudomonas perfectomarinus*. *FEBS Lett.* **148,** 107–112 (1982).
21. Coyle, C. L., Zumft, W. G., Kroneck, P. M. H., Körner, H. & Jakob, W. Nitrous oxide reductase from denitrifying *Pseudomonas perfectomarina*—purification and properties of a novel multicopper enzyme. *Eur. J. Biochem.* **153,** 459–467 (1985).
22. Riester, J., Zumft, W. G. & Kroneck, P. M. H. Nitrous oxide reductase from *Pseudomonas stutzeri*—redox properties and spectroscopic characterization of different forms of the multicopper enzyme. *Eur. J. Biochem.* **178,** 751–762 (1989).
23. Fujita, K., Chan, J. M., Bollinger, J. A., Alvarez, M. L. & Dooley, D. M. Anaerobic purification, characterization and preliminary mechanistic study of recombinant nitrous oxide reductase from *Achromobacter cycloclastes*. *J. Inorg. Biochem.* **101,** 1836–1844 (2007).
24. Dell'Acqua, S. *et al.* A new $Cu_Z$ active form in the catalytic reduction of $N_2$O by nitrous oxide reductase from *Pseudomonas nautica*. *J. Biol. Inorg. Chem.* **15,** 967–976 (2010).
25. Kroneck, P. M. H., Antholine, W. A., Riester, J. & Zumft, W. G. The cupric site in nitrous oxide reductase contains a mixed-valence [Cu(II),Cu(I)] binuclear center: a multifrequency electron-paramagnetic resonance investigation. *FEBS Lett.* **242,** 70–74 (1988).
26. Brown, K. *et al.* Revisiting the catalytic $Cu_Z$ cluster of nitrous oxide ($N_2$O) reductase: evidence of a bridging inorganic sulfur. *J. Biol. Chem.* **275,** 41133–41136 (2000).
27. Shriver, D. F. & Atkins, P. W. *Inorganic Chemistry* 3rd edn, 227–236 (Oxford Univ. Press, 2001).
28. Solomon, E. I. Spectroscopic methods in bioinorganic chemistry: blue to green to red copper sites. *Inorg. Chem.* **45,** 8012–8025 (2006).
29. Paulat, F. *et al.* Spectroscopic properties and electronic structure of pentammineruthenium(II) dinitrogen oxide and corresponding nitrosyl complexes: binding mode of $N_2$O and reactivity. *Inorg. Chem.* **43,** 6979–6994 (2004).
30. Pomowski, A., Zumft, W. G., Kroneck, P. M. H. & Einsle, O. Crystallization of purple nitrous oxide reductase from *Pseudomonas stutzeri*. *Acta Crystallogr. F* **66,** 1541–1543 (2010).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** This work was supported by the Deutsche Forschungsgemeinschaft (IRTG 1422) and the Center for Biological Signalling Studies (*bioss*, EXC 294). We thank the beam line staff at the Swiss Light Source for their assistance with data collection, T. Spatzal, K. Dörner and T. Friedrich for recording EPR spectra and F. Neese for discussions.

**Author Contributions** A.P., W.G.Z. and O.E. performed the experiments, A.P. and O.E. built and refined the structural model, P.M.H.K. and O.E. designed the experiments, O.E., P.K. and W.G.Z. wrote the manuscript.

**Author Information** Atomic coordinates for the reported crystal structures have been deposited with the Protein Data Bank under accession codes 3SBP (native1), 3SBQ (native2) and 3SBR ($N_2$O adduct). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to O.E. (einsle@biochemie.uni-freiburg.de).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Protein purification.** *P. stutzeri* $N_2$OR was purified from fresh cells under strict exclusion of dioxygen and purification was carried out without interruption as described previously[22,31]. The protein was kept constantly under the $N_2/H_2$ atmosphere of a glove box at $<1$ p.p.m. of $O_2$. For long-term storage protein was concentrated to 16 mg ml$^{-1}$ and pelleted by dripping into liquid $N_2$.

**Crystallization and formation of substrate complexes.** *P. stutzeri* $N_2$OR was crystallized under exclusion of dioxygen using the sitting-drop vapour diffusion method as described previously[30]. Purple crystals were harvested into a cryoprotective buffer containing reservoir solution with added 10% (v/v) of 2*R*,3*R*-butane diol and flash cooled in liquid nitrogen for data collection at 100 K. Data sets were recorded on beam lines X06SA and X06DA at the Swiss Light Source. The native *P*1 data set was collected on an in-house rotating anode generator (Rigaku MicroMax 007HF) with an imaging plate (mar research mar345dtb). For data collection statistics see Supplementary Table 1.

For the preparation of substrate complexes, crystals of $N_2$OR were transferred to the cryoprotective buffer solution and subsequently pressurized with 15 bar of pure $N_2O$ gas for 15 min using a XeCell pressurization chamber (Oxford Cryosystems). After pressure release the crystals were immediately flash cooled in liquid nitrogen for diffraction data collection.

**Structure solution and refinement.** The structure of *P. stutzeri* $N_2$OR was solved by the molecular replacement method, using the enzyme from *Paracoccus denitrificans* (PDB code 1FWX) as a search model[26]. One homodimer was located in the asymmetric unit of the $P6_5$ crystal form of the enzyme and four dimers in the *P*1 crystal form. The occurrence of the two crystal forms depended on the crystallization temperature, with hexagonal crystals dominating at 298 K and the *P*1 crystal form at 293 K using identical precipitant solutions[30]. Model building was carried out in Coot[32] and the structures were refined with REFMAC5[33]. In order to assess the structural inhomogeneity observed at the $Cu_Z$ site, restraints were generated from the most well-defined copies of the centre and used for the refinement of all sites. The distances shown in Fig. 3b are based on these restraints. Final refinement statistics are summarized in Supplementary Table 1. Structure representations were created using PyMOL[34].


31. Zumft, W. G. & Matsubara, T. A novel kind of multi-copper protein as terminal oxidoreductase of nitrous oxide respiration in *Pseudomonas perfectomarinus*. *FEBS Lett.* **148,** 107–112 (1982).
32. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66,** 486–501 (2010).
33. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53,** 240–255 (1997).
34. DeLano, W. L. The PyMOL molecular graphics system ⟨http://www.pymol.org⟩ (2002).


©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10339

# Different patterns of peripheral migration by memory CD4$^+$ and CD8$^+$ T cells

Thomas Gebhardt[1], Paul G. Whitney[1], Ali Zaid[1], Laura K. Mackay[1], Andrew G. Brooks[1], William R. Heath[1], Francis R. Carbone[1] & Scott N. Mueller[1]

**Infections localized to peripheral tissues such as the skin result in the priming of T-cell responses that act to control pathogens. Activated T cells undergo migrational imprinting within the draining lymph nodes[1], resulting in memory T cells that provide local and systemic protection[2]. Combinations of migrating and resident memory T cells have been implicated in long-term peripheral immunity, especially at the surfaces that form pathogen entry points into the body[3]. However, T-cell immunity consists of separate CD4$^+$ helper T cells and CD8$^+$ killer T cells, with distinct effector and memory programming requirements[4]. Whether these subsets also differ in their ability to form a migrating pool involved in peripheral immunosurveillance or a separate resident population responsible for local infection control has not been explored. Here, using mice, we show key differences in the migration and tissue localization of memory CD4$^+$ and CD8$^+$ T cells following infection of the skin by herpes simplex virus. On resolution of infection, the skin contained two distinct virus-specific memory subsets; a slow-moving population of sequestered CD8$^+$ T cells that were resident in the epidermis and confined largely to the original site of infection, and a dynamic population of CD4$^+$ T cells that trafficked rapidly through the dermis as part of a wider recirculation pattern. Unique homing-molecule expression by recirculating CD4$^+$ T effector-memory cells mirrored their preferential skin-migratory capacity. Overall, these results identify a complexity in memory T-cell migration, illuminating previously unappreciated differences between the CD4$^+$ and CD8$^+$ subsets.**

Intravital microscopy has provided novel insights into the migration of naive and effector T cells in lymphoid and non-lymphoid compartments, although there has been no concurrent investigation of CD4$^+$ and CD8$^+$ T-cell responses in the context of infection[5,6]. We wanted to examine long-term peripheral immunity by visualizing how the localization of helper and killer T cells differs as immunity progresses to the memory phase of the response. To this end, we used skin infection with herpes simplex virus (HSV) because of the accessibility of this tissue to intravital imaging and the availability of TCR transgenic mice specific for viral helper (gDT-II)[7] and killer (gBT-I)[8] T-cell determinants. Transgenic gDT-II (CD4$^+$) and gBT-I (CD8$^+$) cells expressing fluorescent markers (enhanced green fluorescent protein, EGFP, or the red fluorescent protein, DsRed) were transferred into recipient C57BL/6 mice to track virus-specific T cells. Initial experiments using confocal microscopy showed that both gDT-II and gBT-I T cells were intermixed and present in large numbers throughout the skin 8 days after inoculation (Fig. 1a), which correlates with the time when virus is just cleared from this tissue[9]. Subsequently, the memory phase was marked by segregation of the CD4$^+$ T cells to the dermis and CD8$^+$ T cells to the epidermis (Fig. 1a and b), although T cells were also found in association with keratin-rich hair follicles in regions with indistinct boundaries between epithelium and dermis. Differential expression of the $\alpha_E$ integrin subunit CD103 (an adhesion molecule implicated in persistence of intraepithelial T cells[10,11]) on the skin-infiltrating T cells was consistent with differences in localization by the subsets, with the strongest CD103 staining on the CD8$^+$ cells (Fig. 1c).

Differential localization of helper and killer T cells has previously been noted in human skin[12,13] and other tissues, such as gut and genital mucosa[11,14]. Intravaginal HSV infection resulted in localization (after clearance of the infection) of the majority of the virus-specific helper T cells to the lamina propria and of killer T cells to the epithelium (Supplementary Fig. 2), similar to the separation observed after skin infection (Fig. 1a).

To investigate the dynamics associated with the differential T-cell localization, intravital two-photon microscopy was performed at various times after infection. As observed by static imaging, we found a broad distribution of gDT-II and gBT-I T cells throughout the skin during the effector phase of the response (Fig. 1d and Supplementary Movies 1 and 2). This was followed by retreat of the CD8$^+$ T cells to the epidermis and a largely dermal localization for the CD4$^+$ T cells during memory (Fig. 1d and Supplementary Movies 3–5). During the effector phase, movement of the CD4$^+$ and CD8$^+$ T cells through the dermis occurred with similar mean velocities, and was marked by polarized cell shapes and rapid changes in direction (Fig. 1e and Supplementary Fig. 3a and b). Interestingly, epidermal T cells were significantly less motile than those in the dermis, with a lower mean velocity and higher arrest coefficients during this period. This migrational difference between dermal- and epidermal-associated T cells was even more pronounced during memory, when slow-moving epidermal killer T cells of dendritic morphology were observed above the more rapidly trafficking dermal helper T cells (Fig. 1e–g, Supplementary Fig. 3a and Supplementary Movies 3–5). Such findings suggest that the relative motility of cells in the skin was determined by their localization to epidermis or dermis, and was only indirectly subset related.

The distinct localization and motility of the helper and killer T cells probably reflected differences in trafficking capabilities. To test whether HSV-specific helper T cells could preferentially leave the tissue, we cultured explanted skin that had resolved infection and contained virus-specific memory CD4$^+$ and CD8$^+$ T cells. As shown in Fig. 2a and Supplementary Fig. 4, almost half of all helper T cells left the skin during the culture period, compared to less than 5% of killer T cells, suggesting that the latter were probably sequestered from the circulation. To demonstrate such sequestration, we contrived a situation where cells with access to the bloodstream were susceptible to depletion. Adoptive transfer of virus-specific T cells of male origin into female mice results in an immune response directed to the male minor transplantation antigens that eliminates circulating cells bearing male antigens. We surmised that T cells in the epidermis of the skin might be independent from the circulation, and thus largely resistant to rejection.

Figure 2b and Supplementary Fig. 5a and b show that the male-origin virus-specific helper and killer T cells underwent expansion after skin infection with HSV, and then both were lost from the circulation at around 12 days after infection, consistent with anti-male transplant rejection. This rejection was not seen with female-into-female transfer

[1]Department of Microbiology and Immunology, The University of Melbourne, Melbourne, Victoria 3010, Australia.



©2011 Macmillan Publishers Limited. All rights reserved

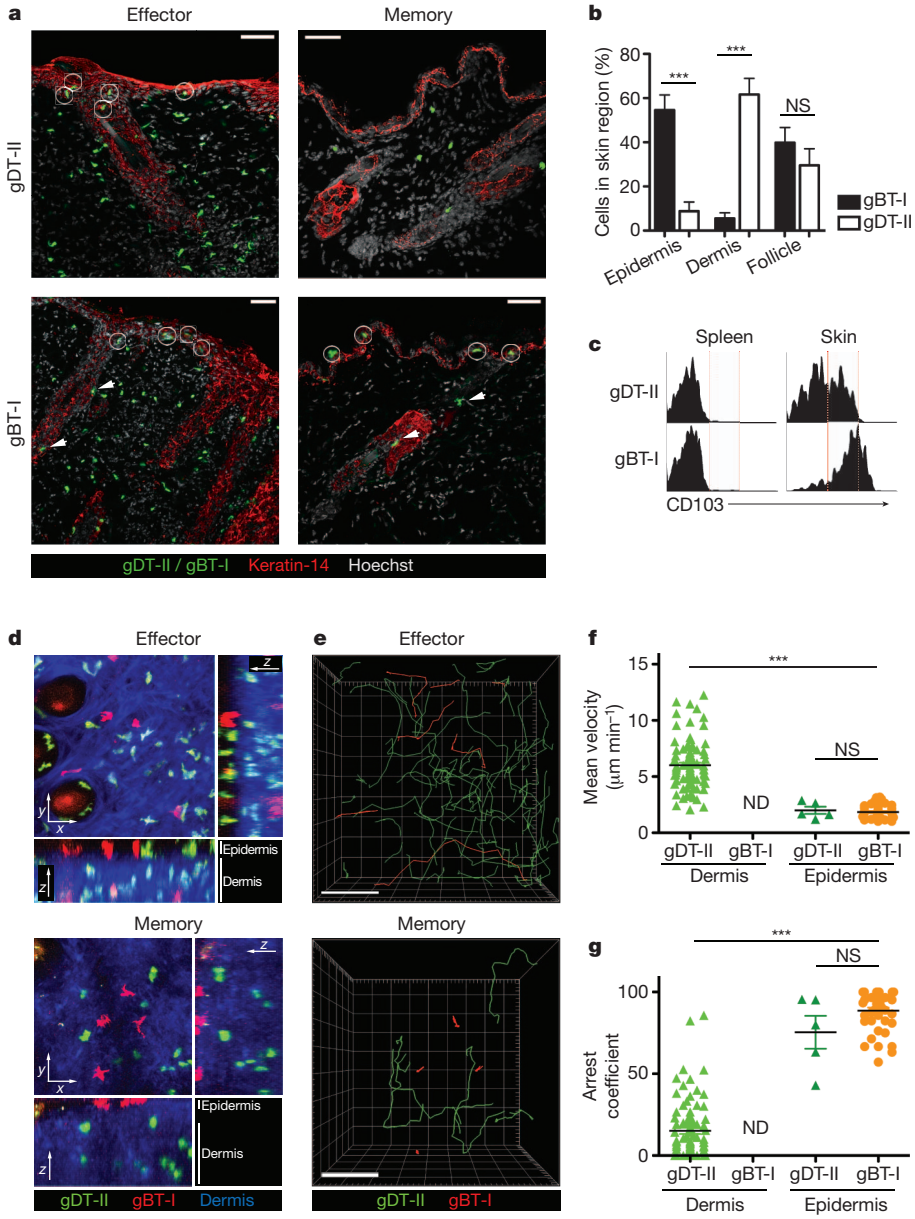



**Figure 1 | Killer and helper memory T cells localize to different areas in the skin and display distinct migration patterns.** **a**, **b**, Immunofluorescence of gBT-I or gDT-II cells (GFP, green) during the effector and memory response after epicutaneous HSV infection (day 8 and 30, respectively). **a**, Sections were stained for keratin-14 (red) and nuclei (Hoechst dye, white). Circles and arrowheads show $GFP^+$ cells in epidermis and around hair follicles, respectively. Scale bars, 50 μm. **b**, The number of gDT-II or gBT-I T cells in different skin regions enumerated 30 days after infection (mean + s.e.m.). $n = 12–14$ sections from 5–6 mice per group. **c**, CD103 expression on spleen and skin T cells 5–6 weeks after infection. **d–g**, Intravital two-photon microscopy of skin gDT-II (GFP, green) and gBT-I (DsRed, red) T cells after HSV infection. **d**, Maximum intensity projections across *x*, *y* and *z* dimensions of skin helper and killer T cells during the effector and memory phases (day 8 and 50 after infection, respectively). Dermis and epidermis was defined by the respective presence or absence of collagen (second harmonic generation, blue). **e**, Cell migration tracks from **d** and Supplementary Movies 1 and 4. **f**, **g**, Skin T-cell mean velocity (**f**) and arrest coefficients (**g**) during memory (day 50 after infection). One representative experiment containing five individual movies from two mice is shown. Points represent individual cells and the lines show the mean ± s.e.m. NS, not statistically significant; ND, not detected; $***P < 0.0001$.

controls (Supplementary Fig. 5c). In contrast to the elimination of both $CD4^+$ and $CD8^+$ T cells from the blood, there was a preferential survival of the male-origin skin-infiltrating killer T cells 4–5 weeks after infecting female recipients (Fig. 2c and d). Remarkably, these $CD8^+$ cells persisted in the absence of their circulating counterparts for at least 10 weeks after infection (Supplementary Fig. 5d and e), demonstrating that they were independent of the circulating memory pool. As the immune privilege responsible for the $CD8^+$ T-cell survival is probably confined to the epidermis, these experiments do not formally exclude the existence of some long-term dermal-resident $CD4^+$ T cells. However, the results in Fig. 1 and 2 suggest that as the infection subsides and the response enters the memory phase, convalescent skin contains two generally distinct virus-specific T-cell subsets; a trafficking population of helper T cells within the dermis largely in equilibrium with the circulation, and a epithelial-localized killer T-cell population that is of limited motility and isolated from the circulating pool.

It has been reported for both sheep and humans that the cells returning to the blood via the afferent lymph show an intriguing bias towards the $CD4^+$ T-cell subset[15,16], suggesting differences in recirculation for the distinct T-cell subsets. Consistent with this, areas of skin remote from the original site of HSV inoculation contained virus-specific $CD4^+$ T cells of high motility after infection had subsided, but strikingly, very few memory $CD8^+$ T cells (Supplementary Fig. 6 and Supplementary Movie 6). We infused whole T-cell populations into uninfected recipients to show that the $CD4^+CD44^+$ memory helper subset dominated entry into the skin and vagina, whereas memory $CD8^+$ T cells were more abundant in the spleen (Fig. 2e), demonstrating preferential helper T-cell migration through these tissues. In line with this, $CD4^+$ T cells also dominated the skin and vagina in non-transfer, uninfected mice (Supplementary Fig. 7).

The distinct behaviour of the peripheral $CD4^+$ and $CD8^+$ T cells suggested that their circulating counterparts might differ in expression of tissue-homing molecules. We examined expression of E-selectin ligands (ESL) and P-selectin ligands (ESL) as E- and P-selectins are key adhesion molecules expressed in the skin endothelium[17], and used splenic populations as surrogates for the circulating pool. In addition to epicutaneous infection, we included subcutaneous injection and intranasal inoculation to fully explore the potential for differential migrational programming of the two T-cell subsets. We note that all three modalities gave roughly equal levels of T-cell priming (Supplementary Fig. 8). Focusing first on the helper T cells, ESL and PSL were strongly upregulated on the effector cells after both forms of skin inoculation (Supplementary Fig. 9a–d), and this expression was generally maintained during memory (Fig. 3a–d). Importantly,



©2011 Macmillan Publishers Limited. All rights reserved

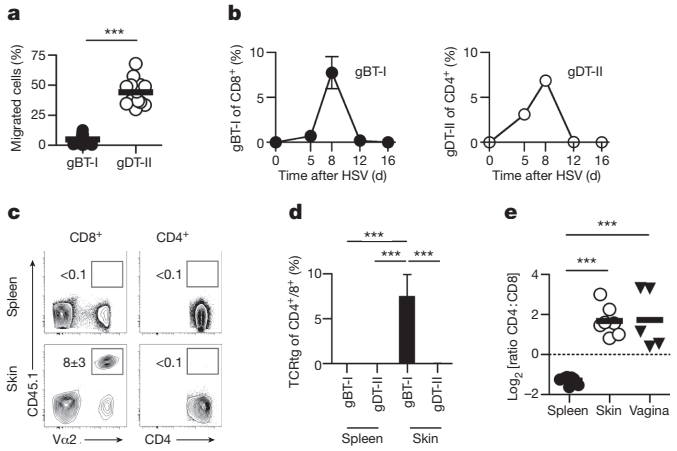



**Figure 2 | Migratory differences between helper and killer memory T cells.** **a**, Explanted skin containing either gBT-I or gDT-II CD45.1$^+$ T cells 7–12 weeks after epicutaneous HSV infection was cultured overnight. Shown is the proportion of cells leaving the skin (migrated cells) compared with cells remaining in the tissue. Data pooled from three independent experiments; $n = 14–15$. **b–d**, Analysis of male gBT-I and gDT-II CD45.1$^+$ cells after transfer into female recipients and epicutaneous HSV infection. **b**, Frequency of HSV-specific effector gBT-I and gDT-II T cells (among total CD8$^+$ or CD4$^+$ T cells, respectively) in the peripheral blood. Results are mean ± s.e.m.; $n = 10$. **c**, **d**, Analysis of gBT-I and gDT-II CD45.1$^+$ cells isolated from spleen and skin 4–5 weeks after infection. **d**, Frequency of HSV-specific memory gBT-I or gDT-II T cells (TCRtg) among total CD8$^+$ or CD4$^+$ T cells, respectively. Results are mean + s.e.m.; $n = 23$. **e**, Ratio of donor CD44$^{hi}$ CD4$^+$ and CD8$^+$ memory T cells in the indicated tissues 6–7 days after transfer of T cells from HSV memory mice into naive recipient mice. Symbols represent individual mice; $n = 5–7$. $^{***}P < 0.001$.

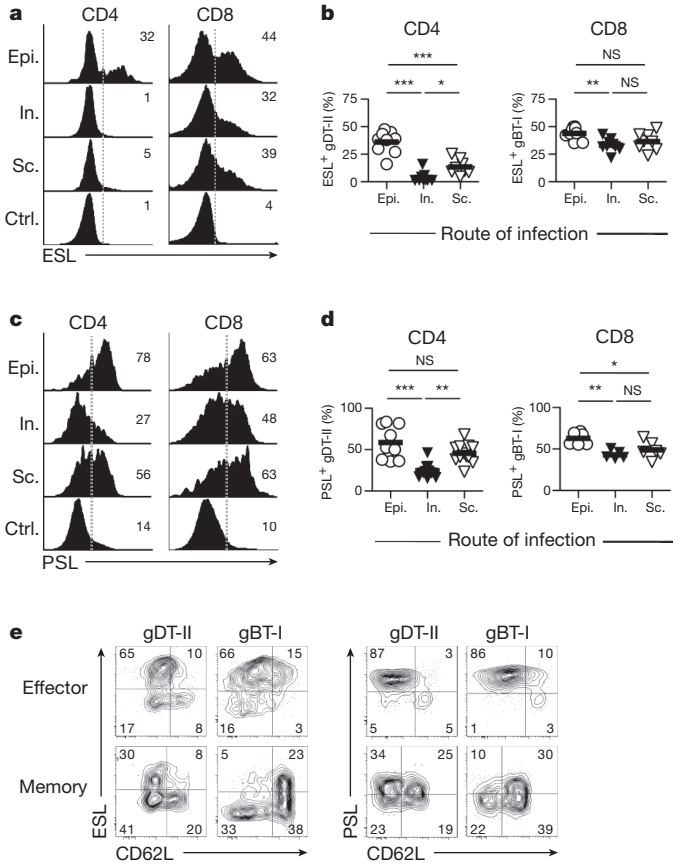



**Figure 3 | Imprinting of ESL and PSL expression in memory T cells.** **a–d**, Analysis of ESL (**a**, **b**) and PSL (**c**, **d**) expression by HSV-specific helper and killer T cells isolated from the spleen 4–8 weeks following adoptive transfer and HSV infection via different routes. Control histograms are gated on total CD8$^+$ cells from naive mice. Epi., epicutaneous; In., intranasal; Sc., subcutaneous; Ctrl., uninfected; $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Numbers on histograms depict percentage of positive cells. Results from 2–4 independent experiments; $n = 8–12$. **e**, CD62L and ESL or PSL co-expression by effector and memory gDT-II and gBT-I CD45.1$^+$ T cells isolated from spleen 6 days or 2–3 months after epicutaneous HSV infection. Numbers represent percentage of events in respective quadrants.

intranasal inoculation induced little ESL or PSL upregulation on the CD4$^+$ effector T cells (Supplementary Fig. 9a–d), and the resultant memory cells were largely ESL/PSL negative (Fig. 3a–d).

In contrast to the CD4$^+$ T cells, all modes of peripheral inoculation drove substantial levels of ESL and especially PSL expression on the CD8$^+$ killer T cells during both the effector (Supplementary Fig. 9a–d) and memory phase of the response (Fig. 3a–d and Supplementary Fig. 9e). At face value, the persisting ESL/PSL levels suggested that all infection modalities generated an equivalent memory killer T-cell population with skin-recirculating capacity. However, closer examination showed that homing molecule expression was largely restricted to CD8$^+$ T cells of the central memory phenotype ($T_{CM}$), calling into question their peripheral surveillance potency (Fig. 3e). Circulating memory T cells are classified into two major subsets: a CD62L$^{lo}$ T effector-memory ($T_{EM}$) population believed to be responsible for tissue surveillance and able to mount rapid response to antigen challenge, and a CD62L$^{hi}$ $T_{CM}$ subset that largely recirculates between the secondary lymphoid organs[18,19]. Whereas there was a marked reduction in homing molecule expression on the CD62L$^{lo}$ $T_{EM}$ subset of CD8$^+$ T cells, CD62L$^{lo}$ $T_{EM}$ CD4$^+$ T cells maintained ESL and PSL expression during the memory phase of the response (Fig. 3e).

Given the preferential expression of ESL/PSL by the circulating CD4$^+$ $T_{EM}$ cells, we ascertained whether this correlated with a dominant role for helper T cells in this form of skin infection. Inhibition of HSV replication was most prominent in mice that were previously infected via the skin either by epicutaneous or subcutaneous means (Fig. 4a); these were the two modalities that programmed long-term ESL and PSL expression in the virus-specific helper $T_{EM}$ subset. Antibody depletion of CD4$^+$ T cells, but not of CD8$^+$ T cells, abrogated protection seen early after HSV infection (Fig. 4b). The assessment of T-cell-mediated protection (Fig. 4a and b) was done using antibody-deficient μMT mice and involved regions of skin not involved in the primary inoculation. This was necessary because antibody can be protective and previously infected skin will contain skin-resident CD8$^+$ memory T cells able to control local infection[20]. Also, virus loads were measured at 48 h after infection, as this corresponds to the initial virus burst at the inoculation site[7,21], and so any reduction in titres would represent an inhibition of the first rounds of infection. Although intrinsic differences in CD4$^+$ and CD8$^+$ T-cell control of HSV infection may explain the observed CD4$^+$ T-cell dominance of protection, the results are nonetheless consistent with the preferential skin-homing ability of this memory subset.

Skin infiltration by the circulating memory CD4$^+$ T cells was examined by delivering a non-specific inflammatory stimulus using the contact-sensitizing agent 2,4-dinitro-1-fluorobenzene (DNFB). At 42 h after DNFB treatment, HSV-specific helper T cells showed preferential accumulation over killer T cells (Supplementary Fig. 10a). As expected, the recruited CD4$^+$ T cells were predominantly CD62L$^{lo}$ $T_{EM}$ cells, whereas the small component of recruited memory CD8$^+$ T cells contained a considerable proportion of CD62L$^{hi}$ $T_{CM}$ (Supplementary Fig. 10b). Circulating memory helper T cells (Fig. 4c) were most efficiently recruited by DNFB-mediated inflammation in mice primed by epicutaneous infection (Fig. 4d and e), and this biased recruitment was also seen in non-inflamed skin (Fig. 4f). Efficient recruitment of the virus-specific T cells was blocked by anti-E-selectin and anti-P-selectin antibodies (Supplementary Fig. 10c). Together, the results suggest that early protection in previously non-involved regions of skin was largely under helper T-cell control, which matched the migrational phenotype evident for this subset.

Our study reveals differential localization and migration of populations of memory CD4$^+$ and CD8$^+$ T cells after cutaneous HSV infection



©2011 Macmillan Publishers Limited. All rights reserved

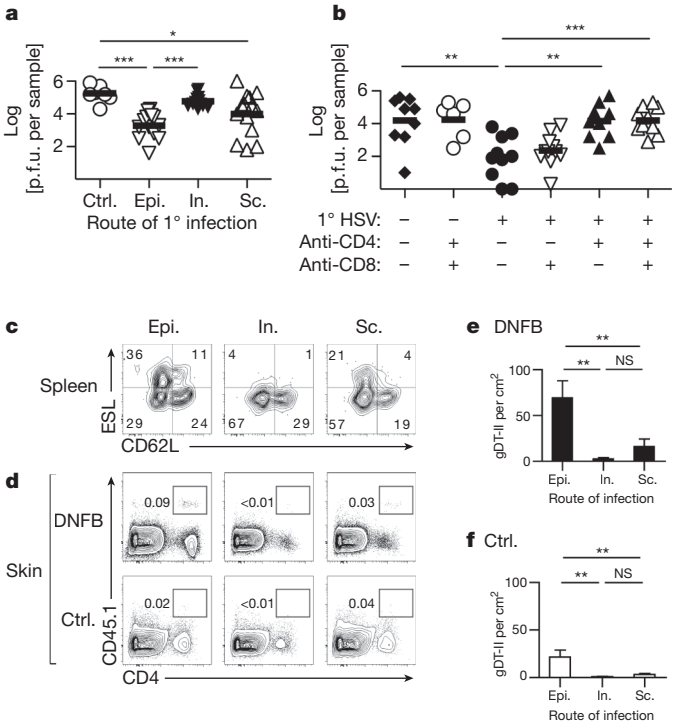



**Figure 4 | Migrational imprinting mirrors immunosurveillance function by memory helper T cells. a**, HSV load in skin 2 days after epicutaneous challenge of μMT mice primed via different routes of inoculation (as in Fig. 3) 4–6 weeks earlier. Controls (Ctrl.) are non-immunized mice. $n = 9$–14. **b**, HSV load in contralateral skin 2 days after epicutaneous inoculation of μMT mice primed 4–5 weeks earlier and treated with anti-CD4 and/or anti-CD8 antibodies before reinfection. $n = 6$–11. **c**–**f**, Analysis of HSV-specific memory T helper cells isolated from spleen and inflamed (DNFB) or control (Ctrl.) skin. DNFB was applied to skin of mice containing gDT-II CD45.1$^+$ memory CD4$^+$ T cells after HSV inoculation via different routes (as in Fig. 3) 9–12 weeks earlier. **c**, Phenotype of memory gDT-II CD45.1$^+$ T cells isolated from spleen. Numbers represent percentage of events in respective quadrants. **d**–**f**, Detection and enumeration of memory gDT-II CD45.1$^+$ T cells in skin 42 h after initiation of skin inflammation expressed as mean + s.e.m. $n = 9$–11; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

(Supplementary Fig. 1). It remains unclear whether the phenomenon extends to other infections, although preferential recirculation of CD4$^+$ T cells has been shown for normal skin of sheep and humans[15,16]. In addition, sequestration of CD8$^+$ memory T cells within extra-lymphoid compartments, with limited input from the wider circulation, has been reported with other tissue and virus combinations[22–24]. Overall, the previously unappreciated migration-linked T-cell subset specialization shown here has the potential to affect a range of peripheral immune phenomena, at least in barrier tissues such as skin and mucosa.

## METHODS SUMMARY

Epicutaneous infection by scarification was done using the HSV-1 KOS strain as described elsewhere[9]. For T-cell isolation, mice were perfused with saline before removal of 1–2 cm$^2$ of skin, which was incubated for 90 min at 37 °C with collagenase (3 mg ml$^{-1}$) and DNase (5 μg ml$^{-1}$). For some experiments, the skin was first incubated for 90 min at 37 °C with Dispase (2.5 mg ml$^{-1}$) and DNase (5 μg ml$^{-1}$), followed by mechanical separation of epidermis and dermis. For intravital two-photon microscopy, mice were anaesthetized, depilated and two parallel incisions were made approximately 15 mm apart along the flank. The skin was separated from the peritoneum, and adhered to an 18-mm-wide piece of 1-mm stainless steel inserted to form a stable raised platform, attached to a custom-made imaging platform maintained at 35 °C. Images were acquired with an upright LSM710 NLO multiphoton microscope (Carl Zeiss Microimaging) with a ×20/1.0 NA water immersion objective. For four-dimensional data sets, three-dimensional stacks were captured every 1 min for 30–120 min. Raw imaging data were processed with Imaris 7.1 (Biplane).

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 17 September 2010; accepted 28 June 2011.**
**Published online 14 August 2011.**

1. Sigmundsdottir, H. & Butcher, E. C. Environmental cues, dendritic cells and the programming of tissue-selective lymphocyte trafficking. *Nature Immunol.* **9,** 981–987 (2008).
2. Kaech, S. M. & Wherry, E. J. Heterogeneity and cell-fate decisions in effector and memory CD8$^+$ T cell differentiation during viral infection. *Immunity* **27,** 393–405 (2007).
3. Woodland, D. L. & Kohlmeier, J. E. Migration, maintenance and recall of memory T cells in peripheral tissues. *Nature Rev. Immunol.* **9,** 153–161 (2009).
4. Seder, R. A. & Ahmed, R. Similarities and differences in CD4$^+$ and CD8$^+$ effector and memory T cell generation. *Nature Immunol.* **4,** 835–842 (2003).
5. Bousso, P. T-cell activation by dendritic cells in the lymph node: lessons from the movies. *Nature Rev. Immunol.* **8,** 675–684 (2008).
6. Mueller, S. N. & Hickman, H. D. *In vivo* imaging of the T cell response to infection. *Curr. Opin. Immunol.* **22,** 293–298 (2010).
7. Bedoui, S. *et al.* Cross-presentation of viral and self antigens by skin-derived CD103$^+$ dendritic cells. *Nature Immunol.* **10,** 488–495 (2009).
8. Mueller, S. N., Heath, W., McLain, J. D., Carbone, F. R. & Jones, C. M. Characterization of two TCR transgenic mouse lines specific for herpes simplex virus. *Immunol. Cell Biol.* **80,** 156–163 (2002).
9. van Lint, A. *et al.* Herpes simplex virus-specific CD8$^+$ T cells can clear established lytic infections from skin and nerves and can partially limit the early spread of virus after cutaneous inoculation. *J. Immunol.* **172,** 392–397 (2004).
10. Cepek, K. L. *et al.* Adhesion between epithelial cells and T lymphocytes mediated by E-cadherin and the $\alpha^E\beta_7$ integrin. *Nature* **372,** 190–193 (1994).
11. Schon, M. P. *et al.* Mucosal T lymphocyte numbers are selectively reduced in integrin $\alpha_E$ (CD103)-deficient mice. *J. Immunol.* **162,** 6641–6649 (1999).
12. Nestle, F. O., Di Meglio, P., Qin, J. Z. & Nickoloff, B. J. Skin immune sentinels in health and disease. *Nature Rev. Immunol.* **9,** 679–691 (2009).
13. Zhu, J. *et al.* Persistence of HIV-1 receptor-positive cells after HSV-2 reactivation is a potential mechanism for increased HIV-1 acquisition. *Nature Med.* **15,** 886–892 (2009).
14. Miller, C. J., McChesney, M. & Moore, P. F. Langerhans cells, macrophages and lymphocyte subsets in the cervix and vagina of rhesus macaques. *Lab. Invest.* **67,** 628–634 (1992).
15. Mackay, C. R., Marston, W. L. & Dudler, L. Naive and memory T cells show distinct pathways of lymphocyte recirculation. *J. Exp. Med.* **171,** 801–817 (1990).
16. Yawalkar, N., Hunger, R. E., Pichler, W. J., Braathen, L. R. & Brand, C. U. Human afferent lymph from normal skin contains an increased number of mainly memory/effector CD4$^+$ T cells expressing activation, adhesion and co-stimulatory molecules. *Eur. J. Immunol.* **30,** 491–497 (2000).
17. Ley, K., Laudanna, C., Cybulsky, M. I. & Nourshargh, S. Getting to the site of inflammation: the leukocyte adhesion cascade updated. *Nature Rev. Immunol.* **7,** 678–689 (2007).
18. Masopust, D., Vezys, V., Marzo, A. L. & Lefrancois, L. Preferential localization of effector memory cells in nonlymphoid tissue. *Science* **291,** 2413–2417 (2001).
19. Sallusto, F., Lenig, D., Forster, R., Lipp, M. & Lanzavecchia, A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature* **401,** 708–712 (1999).
20. Gebhardt, T. *et al.* Memory T cells in nonlymphoid tissue that provide enhanced local immunity during infection with herpes simplex virus. *Nature Immunol.* **10,** 524–530 (2009).
21. Wakim, L. M., Jones, C. M., Gebhardt, T., Preston, C. M. & Carbone, F. R. CD8$^+$ T-cell attenuation of cutaneous herpes simplex virus infection reduces the average viral copy number of the ensuing latent infection. *Immunol. Cell Biol.* **86,** 666–675 (2008).
22. Masopust, D. *et al.* Dynamic T cell migration program provides resident memory within intestinal epithelium. *J. Exp. Med.* **207,** 553–564 (2010).
23. Wakim, L. M., Woodward-Davis, A. & Bevan, M. J. Memory T cells persisting within the brain after local infection show functional adaptations to their tissue of residence. *Proc. Natl Acad. Sci. USA* **107,** 17872–17879 (2010).
24. Klonowski, K. D. *et al.* Dynamics of blood-borne CD8 memory T cell migration *in vivo*. *Immunity* **20,** 551–562 (2004).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank members of the Carbone and Heath laboratories for discussions. This work was supported by the Australian NHMRC and ARC.

**Author Contributions** T.G., W.R.H., F.R.C. and S.N.M. designed the study. T.G., A.Z., L.K.M. and S.N.M. performed experiments. T.G., A.Z. and S.N.M. collected and analysed data. P.G.W. and A.G.B. provided reagents and gave conceptual advice. F.R.C. wrote the manuscript with T.G., W.R.H. and S.N.M. The research program was led by F.R.C. and W.R.H. who are joint senior authors.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.N.M. (smue@unimelb.edu.au) or T.G. (gebhardt@unimelb.edu.au).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Mice.** C57BL/6, B6.SJL-*PtprcaPep3b*/BoyJ (B6.CD45.1), gBT-I[8], gDT-II[7], gBT-I × B6.CD45.1 (gBT-I.CD45.1), gBT-I.EGFP, gBT-I.DsRed, gDT-II × B6.CD45.1 (gDT-II.CD45.1), gDT-II.EGFP and μMT mice were bred in the Department of Microbiology and Immunology at The University of Melbourne. gBT-I and gDT-II are T-cell receptor-transgenic mice recognizing the HSV-1 glycoprotein B-derived epitope $gB_{498–505}$ and the glycoprotein D-derived epitope $gD_{290–302}$, respectively. B6.CD45.1 mice express the congenic marker CD45.1, in contrast to C57BL/6 that express CD45.2. gBT-I.CD45.1 and gDT-II.CD45.1 are $F_1$ generation offspring expressing both CD45.1 and CD45.2. Animal experiments were approved by The University of Melbourne Animal Ethics Committee.

**Virus infections.** Epicutaneous infection by scarification was done using $1 \times 10^6$ plaque forming units (p.f.u.) of the HSV-1 KOS strain as described in detail elsewhere[9]. Briefly, mice were anaesthetized by i.p. injection of a 1:1 mixture of ketamine and Xylazil (10 μl $g^{-1}$ of body weight) solution in saline. Hair was removed from the left flank of each mouse with clippers and depilation cream (Veet, Reckitt Benckiser) and a small area of skin (2–4 $mm^2$) corresponding with the top of the spleen was lightly abraded before inoculation of virus in a 10 μl volume. The infection site was then covered with a piece of OpSite Flexigrid (Smith & Nephew) and secured with Micropore tape and Transpore tape (3M Health Care) for 48 h following infection. This method of inoculation results in zosteriform spread of virus throughout the dermatome with complete virus clearance by day 8 after infection[9] and extinguishing of antigen presentation within the lymph nodes by 3–4 weeks after infection[25]. Lung infection ($1 \times 10^6$ p.f.u.) was performed by intranasal inoculation and subcutaneous infection by injection into the left and right foot hocks ($4 \times 10^5$ p.f.u. each). Intravaginal infection was performed on mice treated with progesterone (Depo Provera, 2 mg per mouse) by subcutaneous injection in the scruff of the neck, 5–7 days earlier. For infection, mice were intravaginally swabbed with calcium alginate swabs and inoculated with $1 \times 10^6$ p.f.u. of HSV in 10 μl phosphate buffered saline (PBS) using a blunt pipette tip. Rechallenge infections, done in μMT mice to circumvent interference by neutralizing antibodies, involved scarification using the thymidine kinase-deficient KOS.CreTK$^-$ strain ($1 \times 10^6$ p.f.u.). Viral titres were analysed using standard p.f.u. assay on Vero cells as described elsewhere[9].

**Adoptive transfer of T cells and antibody treatments.** T cells were isolated from spleen and/or lymph nodes of transgenic female mice. For selected experiments, cell suspensions were magnetically enriched for $CD8^+$ or $CD4^+$ T cells by negative selection as described[7], yielding purities of >90% for gBT-I and 20–40% for gDT-II cells. For some experiments, gDT-II cells were further positively selected using magnetically labelled anti-CD4 antibodies (Dynabeads mouse CD4, L4T4) and the DetachABead reagent for mouse CD4 (both from Invitrogen Dynal As), yielding purities between 40% and 70%. Cell suspensions corresponding to $1 \times 10^4$ gDT-II ($2.5 \times 10^4$ for imaging) and $5 \times 10^4$ gBT-I cells were intravenously transferred into congenic female recipient mice. For sex-mismatched transfer experiments, cells were isolated from male donor mice and $1 \times 10^5$ of each transgenic T-cell population was transferred either together or separately into recipient female mice. For experiments analysing the migration of endogenous memory T cells, spleen cells were isolated from C57BL/6 mice 1–7 months after HSV flank infection, magnetically enriched for T cells and transferred into naive B6.CD45.1 recipient mice ($3–7 \times 10^7$ total cells per mouse). Donor cells were analysed in spleen, skin and vagina at least 6 days after transfer. Depletion of $CD4^+$ and/or $CD8^+$ cells was achieved by treating mice with monoclonal antibodies, GK1.5 (0.1 mg per mouse) and/or 2.43 (0.1 mg per mouse) for three consecutive days. These antibodies were prepared by the National Cell Culture Center. Blocking antibodies to CD62E and CD62P (10E9.6, RB40.34; both from BD Pharmingen) were injected intraperitoneally into mice (0.1 mg of each antibody) for three consecutive days before DNFB challenge. The treatment was continued for 2 days (0.2 mg of each antibody) thereafter.

**Immunofluorescence and confocal microscopy.** Skin from the flank of HSV-infected mice (left flank, infected; right flank, uninfected) was collected at various times after infection and fixed in 4% paraformaldehyde and 10% sucrose (w/v in PBS) for 20 min at room temperature (20 °C). Vaginal tissues were collected from mice after intravaginal infection and fixed in 4% paraformaldehyde and 10% sucrose for 10 min at room temperature (20 °C), followed by incubation in 20% sucrose for 20 min. Fixed tissue was embedded in OCT (Tissue Tek IA018; Sakura) and frozen in liquid nitrogen. Tissue sections were performed on a cryostat (Leica CM3050S) at 12–25 μm thickness and air-dried for 16–18 h before staining. Sections were fixed in −20 °C acetone for 5 min, rehydrated with PBS for 15 min then blocked for 10 min (Protein Block X0909; DAKO) at room temperature (20 °C). When using biotinylated antibodies, avidin-biotin blocking (Dako Biotin-Blocking System X0590; DAKO) was applied for 5 min each before primary antibody incubation. Sections were stained with rabbit anti-keratin 14 antibody (clone AF64; Covance), rabbit anti-collagen IV (Abcam), or CD8a-AlexaFluor488 (clone 53-6.7; Caltag) and CD4-biotin (clone L3T4, BD Pharmingen) for 3 h at 4 °C in a semi-humid chamber. Sections were rinsed for 10 min in PBS, and labelled with donkey anti-rabbit AlexaFluor647 (A31573; Invitrogen) or streptavidin-conjugated AlexaFluor555 (S32355; Invitrogen) for 40 min at room temperature (20 °C). Sections were rinsed for 10 min in PBS, incubated with Hoechst nuclear stain (H33258, 1:3,000 vol./vol. in PBS) for 3 min, and then mounted with ProLongGold (P36934; Invitrogen). Images were acquired with a Zeiss LSM710 microscope and processed using Imaris 7.1 software (Bitplane). For calculating T-cell numbers in skin regions (epidermis, dermis and associated with hair follicles) or vaginal tissue (epithelia and lamina propria), the number of T cells in each region was counted in multiple sections from each mouse at 30 days after infection (skin) or 22 days after infection (vagina).

**Intravital two-photon microscopy.** Mice were anaesthetized with isofluorane (Cenvet; 2.5% for induction, ~1.5% for maintenance, vaporized in an 80:20 mixture of oxygen and air) using a Tech 3 vaporizer (Surgivet), and the left flank shaved and depilated using Veet (Reckitt Benckiser). Two parallel incisions were made through the dermis, approximately 15 mm apart, along the flank. The skin was carefully separated from the peritoneum, and an 18-mm-wide piece of 1-mm stainless steel inserted under the dermis. This formed a stable raised platform, attached to a custom made imaging platform. The underside of the dermis was adhered to the steel platform using Vetbond tissue adhesive (3M). Vacuum grease (Dow Corning) was used to attach and seal a glass coverslip onto the epidermis of the skin for imaging. Moist gauze was packed around the incision and covered to prevent dehydration.

Images were acquired with an upright LSM710 NLO multiphoton microscope (Carl Zeiss Microimaging) enclosed in a custom-built environmental chamber (Precision Plastics) that was maintained at 35 °C with heated air. Three external non-descanned PMT detectors in the reflected light path were used to acquire images. Images were acquired with a ×20/1.0 NA water immersion objective. Fluorescence excitation was provided by a Chamelon Vision II Ti:sapphire laser (Coherent), with dispersion correction. EGFP and DsRed were excited at 940 nm. The collagen-rich dermis of the skin was visualized by second harmonic generation, using a short-pass 485 nm filter. Typical voxel dimensions were (0.55–0.7) × (0.55–0.7) × 3 μm. For four-dimensional data sets, three-dimensional stacks were captured every 1 min for 30–120 min.

Raw imaging data were processed with Imaris 7.1 (Biplane). Cell migration was analysed through automatic cell tracking aided by manual corrections. Only tracks that lasted longer than 5 min were analysed. Statistical data were imported into Excel (Microsoft) for calculating cell velocities, displacement over time, relative cell tracks, and arrest coefficients (proportion of time cells were moving at less than 2 μm $min^{-1}$). Data were used to plot graphs in Prism 5 (Graphpad). For the generation of movies, image sequences exported from Imaris were composed in Adobe After Effects CS5.

**Cell isolation and flow cytometry.** Mice were killed by carbon dioxide administration and perfusion performed with 10 ml Hank's buffered salt solution (Media Preparation Unit, The University of Melbourne). Spleens were harvested and disrupted by passage through a wire mesh. Skin tissue (1–2 $cm^2$) was removed after clipping and treatment of flank skin with Veet depilation cream (Reckitt Bencksier), chopped into small fragments and incubated for 90 min at 37 °C in MEM medium containing fetal calf serum (FCS, 2%, Gibco BRL), collagenase (3 mg $ml^{-1}$; Worthington) and DNase (5 μg $ml^{-1}$, Sigma). For some experiments skin samples were first incubated for 90 min at 37 °C in 3 ml PBS containing Dispase (2.5 mg $ml^{-1}$, Roche) and DNase (5 μg $ml^{-1}$) followed by the mechanical separation of epidermis and dermis. Epidermal sheets were subsequently incubated for 30 min in Trypsin/EDTA (0.25%/0.1%, SAFC Biosciences) whereas the remaining skin tissue was chopped into small fragments and incubated for 30 min at 37 °C in collagenase solution. For analysis of vaginal tissues, tissue was isolated, the cervix removed, and chopped into small pieces before incubation in calcium- and magnesium-free HBSS medium containing 1.3 mM EDTA for 30 min (37 °C, shaking at 180 r.p.m.). Tissue pieces were then treated with 1 mg $ml^{-1}$ collagenase 3 (Worthington) and 5 μg $ml^{-1}$ DNase (Sigma) in RPMI supplemented with 1 mM $MgCl_2$ and $CaCl_2$ and 5% FCS for 90 min at 37 °C, shaking at 180 r.p.m. Tissues were pushed through 70 μm nylon mesh and lymphocytes purified on a 44/67% Percoll (GE Healthcare) gradient (centrifuged at 800*g*, 20 °C for 20 min, no brake). All cell suspensions were filtered twice through nylon meshes (70 and 30 μm pore size) before staining for flow cytometric analysis (30 min at 4 °C) or overnight culture in RPMI 1640 medium containing 10% (vol./vol.) FCS and standard supplements[12]. The following antibodies/reagents were purchased from BD Pharmingen: anti-CD8a-APC and anti-CD8a-PE-Cy7 (53-6.7), anti-CD45.1-FITC and anti-CD45.1-PE (A20), anti-CD45.2-FITC (104), anti-CD4-APCCy7 and anti-CD4-PE-Cy7 (GK1.5 or RM4-5), anti-Vα2-PE (B20.1), CD62L-PE (Mel-14), anti-CD16/32 (2.4G2) and streptavidin-FITC. Anti-CD8a-APC eFluor 780 (53-6.7), anti-CD45.1-APC (A20), anti-CD45.2-PE (104) and anti-CD45.2-APC

©2011 Macmillan Publishers Limited. All rights reserved

eFluor 780 (104) and anti-CD103-FITC (2E7) were purchased from eBioscience. E- and P-selectin binding was measured using mouse E- or P-selectin Fc chimaeric proteins (1 $\mu g\,ml^{-1}$ or 0.3 $\mu g\,ml^{-1}$ respectively; R&D Systems) in conjunction with biotin conjugated rabbit anti-human IgG F(c) antibody (Rockland). Dead cells were identified using propidium iodide (PI) staining. Sphero calibration particles (BD Pharmingen) were added to the samples to allow for calculation of cell numbers. Flow cytometric analysis was performed using FACSCanto II flow cytometers and FlowJo software (TreeStar).

**DNFB treatment.** HSV memory or naive control mice were anaesthetized and the back and the flanks were clipped and treated with Veet depilation cream. 25 µl of 0.5% (vol./vol.) DNFB in acetone/oil (4:1) were applied onto a 1 $cm^2$ area of previously non-infected skin. 42 h after treatment, mice were killed for analysis of memory cell migration into treated and non-treated areas of skin.

***Ex vivo* migration assay.** Full thickness skin was excised from previously infected mice as described above. Skin samples were cultured overnight in RPMI 1640 medium containing 10% (vol./vol.) FCS and standard supplements[9]. The next day, cells that had migrated out of explants were harvested from the culture medium and non-migrated cells left in skin tissue were isolated as described above. Cells were enumerated by flow cytometry.

**Statistics.** Comparison of data sets was performed by one-way analysis of variance followed by Tukey post-test or by two-tailed, unpaired Student's *t*-test where appropriate.

25. Stock, A. T., Jones, C. M., Heath, W. R. & Carbone, F. R. Rapid recruitment and activation of $CD8^+$ T cells after herpes simplex virus type 1 skin infection. *Immunol. Cell Biol.* **89,** 143–148 (2011).

©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10341

# Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry

Mathieu Joron[1,2,3], Lise Frezal[1]*, Robert T. Jones[4]*, Nicola L. Chamberlain[4], Siu F. Lee[5], Christoph R. Haag[6], Annabel Whibley[1], Michel Becuwe[2], Simon W. Baxter[7], Laura Ferguson[7], Paul A. Wilkinson[4], Camilo Salazar[8], Claire Davidson[9], Richard Clark[9], Michael A. Quail[9], Helen Beasley[9], Rebecca Glithero[9], Christine Lloyd[9], Sarah Sims[9], Matthew C. Jones[9], Jane Rogers[9], Chris D. Jiggins[7] & Richard H. ffrench-Constant[4]

**Supergenes are tight clusters of loci that facilitate the co-segregation of adaptive variation, providing integrated control of complex adaptive phenotypes[1]. Polymorphic supergenes, in which specific combinations of traits are maintained within a single population, were first described for 'pin' and 'thrum' floral types in *Primula*[1] and *Fagopyrum*[2], but classic examples are also found in insect mimicry[3–5] and snail morphology[6]. Understanding the evolutionary mechanisms that generate these co-adapted gene sets, as well as the mode of limiting the production of unfit recombinant forms, remains a substantial challenge[7–10]. Here we show that individual wing-pattern morphs in the polymorphic mimetic butterfly *Heliconius numata* are associated with different genomic rearrangements at the supergene locus *P*. These rearrangements tighten the genetic linkage between at least two colour-pattern loci that are known to recombine in closely related species[9–11], with complete suppression of recombination being observed in experimental crosses across a 400-kilobase interval containing at least 18 genes. In natural populations, notable patterns of linkage disequilibrium (LD) are observed across the entire *P* region. The resulting divergent haplotype clades and inversion breakpoints are found in complete association with wing-pattern morphs. Our results indicate that allelic combinations at known wing-patterning loci have become locked together in a polymorphic rearrangement at the *P* locus, forming a supergene that acts as a simple switch between complex adaptive phenotypes found in sympatry. These findings highlight how genomic rearrangements can have a central role in the coexistence of adaptive phenotypes involving several genes acting in concert, by locally limiting recombination and gene flow.**

The origin and maintenance of adaptive multi-locus polymorphism in the face of recombination is a long-standing puzzle in evolutionary biology[7,12,13]. In some cases, supergene architecture has evolved with tight linkage that maintains specific combinations of alleles at neighbouring genes[1–6]. A notable illustration is provided by polymorphic mimetic butterflies, in which several discrete forms, each resembling a different model, are maintained in sympatry. Examples include Batesian polymorphism in *Papilio dardanus*[5,14] and *Papilio memnon*[3] and Müllerian polymorphism in the neotropical species *Heliconius numata*[4,15]. In each case, a single supergene locus controls coordinated differences in a complex phenotype which can involve modifications of wing pattern and shape, body colour and perhaps behaviour[3,4]. Mimetic patterns represent sharp fitness peaks corresponding to locally abundant wing patterns, separated by adaptive valleys in which selection acts against recombinant individuals with intermediate, non-mimetic phenotypes[16].

Theoretical debate has centred on the constraints imposed on supergene evolution by genomic organization, specifically whether loci must be tightly linked from the outset or whether the association between elements can be acquired, either gradually or in a single mutational step[7–10,16–19]. Chromosomal rearrangements, which can bring genes into closer physical association and influence local recombination, offer one route through which supergenes may be assembled from more loosely linked components[7,8,17–19]. Although there are many cases of polymorphic inversions associated with adaptive variation[12,20], variation is in most cases geographical, rather than being maintained within populations, and effects on local adaptation are cumulative. In contrast, supergenes are characterized by a Mendelian switch between clearly defined combinations of traits in populations. Here we investigate the genomic organization and population genetics of the *P* supergene in *H. numata*[4,9] and identify a key role for structural variation in strengthening and maintaining allelic associations within the supergene.

In *H. numata*, up to seven sympatric morphs coexist in local populations and each is an accurate mimic of one of several available model species in another butterfly family (Danainae: *Melinaea*)[4,9]. Each morph is controlled by a specific allele at *P* with precise allelic dominance[4,15]

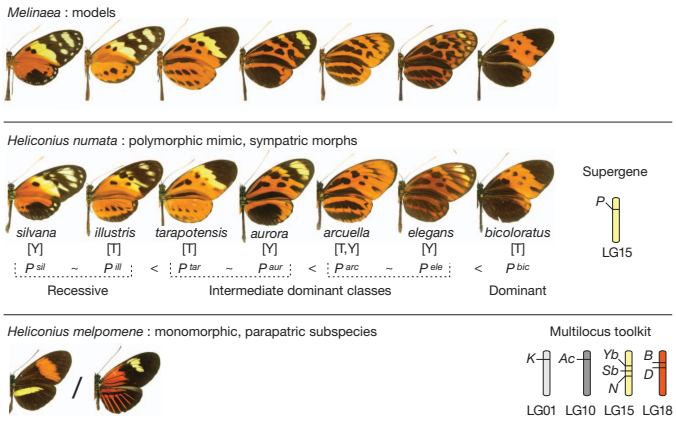



**Figure 1 | Supergene alleles and mimicry polymorphism in *H. numata*.** Polymorphic forms of *H. numata* each mimic different models in the distantly related genus *Melinaea* (Nymphalidae: Danainae). Each form is controlled by a specific allele of the supergene *P*, with increasing dominance shown from left to right[4,9]. Two parapatric regions of northeastern Peru (T, Tarapoto and Andean valleys; Y, Yurimaguas and Amazon lowlands) harbour different mimicry assemblages[15]; dominance (<) is nearly complete between forms within each region, but is incomplete (~) between certain pairs of alleles from parapatric regions. In all other species studied in the genus *Heliconius*, wing pattern is controlled by several large-effect loci on different chromosomes. In *H. melpomene* the *HmYb–HmSb–HmN* complex is situated in the orthologous position to the *H. numata P* supergene[9]. LG, linkage group.

[1]CNRS UMR 7205, Muséum National d'Histoire Naturelle, CP50, 45 Rue Buffon, 75005 Paris, France. [2]Institute of Evolutionary Biology, University of Edinburgh, Ashworth Laboratories, King's Buildings, West Mains Road, Edinburgh EH9 3JT, UK. [3]Institute of Biology, Leiden University, Postbus 9505, 2300 RA Leiden, The Netherlands. [4]Centre for Ecology and Conservation, School of Biosciences, University of Exeter, Cornwall Campus, Penryn, Cornwall TR10 9EZ, UK. [5]Department of Genetics, Bio21 Institute, University of Melbourne, 30 Flemington Road, Parkville, 3010 Victoria, Australia. [6]Department of Biology, Ecology and Evolution, University of Fribourg, Chemin du Musée 10, CH-1700 Fribourg, Switzerland. [7]Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK. [8]Smithsonian Tropical Research Institute, NAOS island, Causeway Amador, Panamá, República de Panamá. [9]The Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1HH, UK.
*These authors contributed equally to this work.



©2011 Macmillan Publishers Limited. All rights reserved

(Fig. 1). Rare non-mimetic individuals that combine pattern elements from different morphs are observed at a frequency of <0.7% in natural polymorphic populations[4,15]. These individuals, presumed to be recombinants, confirm the existence of several functional elements in the locus. The genomic position of *P* was previously shown to correspond to a cluster of three loci, *HmN*, *HmYb* and *HmSb*, on linkage group 15 in the closely related species *Heliconius melpomene*[9], in which they control distinct wing-pattern elements in geographical races. Other unlinked wing-pattern loci in related species do not have large effects on *H. numata* mimicry[9,21].

By fine-scale linkage mapping and positional cloning, we identified a chromosomal interval of about 400 kilobases (kb) containing the *H. numata P* supergene, defined by the absence of crossing over in 366 progeny from six broods (Fig. 2a). The orthologous region in *H. melpomene* shows notable rates of recombination and contains two distinct colour-pattern loci, *HmYb* and *HmSb*, ~0.9 centimorgans (cM) apart[11]. Although genome-wide estimates of recombination rate in *Heliconius* are not available, markers adjacent to *P*, markers elsewhere on linkage group 15, and markers in other linkage groups (Supplementary Table 1)[9] all show significantly higher rates of crossing over (chi-squared test of independence, $P = 3.8 \times 10^{-6}$), consistent with severe suppression of recombination at *P* in *H. numata*.

To examine the gene content and genomic organization of the supergene, we sequenced an approximately one-megabase (Mb) region centred on *P* by screening a bacterial artificial chromosome (BAC) library prepared from a mixture of individuals from a single polymorphic population of *H. numata* (Supplementary Table 2). Colinearity and gene content was generally conserved with *H. melpomene*[11], but two BACs overlapping the *P* interval showed distinct gene orders, both different from the order seen in *H. melpomene* (Figs 2a, 3 and Supplementary Fig. 1). A third gene order, corresponding to the *H. melpomene* reference, was detected by PCR (see below). Gene orientation near the breakpoints on the *H. numata* clones indicated a minimum of two rearrangement events compared to the reference, putatively involving 31 genes: a 400-kilobase (kb) segment containing 18 genes from *HN00023* to *HN00040* (breakpoint BP1), and a 180-kb segment containing 13 genes from *HN00041* to *HN00053* (breakpoint BP2) (Fig. 3a and Supplementary Fig. 1). The altered gene orders at *P* contrast with the colinearity of flanking sequences between *H. numata* and *H. melpomene* (Supplementary Fig. 2 and Supplementary Tables 3 and 4), as well as between more distantly related *Heliconius* species[22]. Furthermore, *H. melpomene* and the silk moth *Bombyx mori*, separated by about 100 million years of evolution, share a generally conserved gene order across this region[11,23]. This supports the hypothesis that the

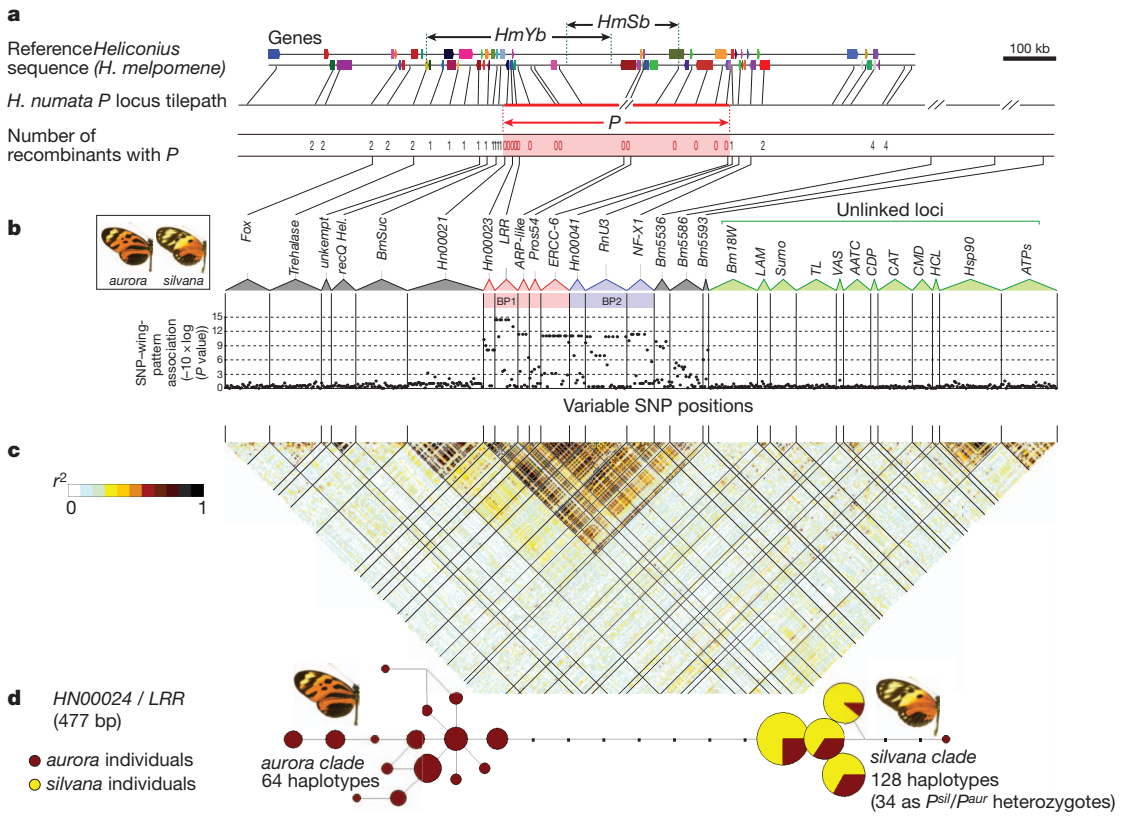



**Figure 2 | Fine-scale mapping and nucleotide variation at the *P* supergene in *H. numata*.** **a**, Fine-scale linkage mapping of the supergene *P* (indicated by red arrows) to the interval bounded by genes *HN00020* and *HN00041*. Recombinants were observed in crosses totalling 366 individuals. Blue arrows indicate the position of the recombining loci *HmYb* and *HmSb* in *H. melpomene*[11]. *HmN*, which is known to be part of this cluster of loci, is not fine-mapped in *H. melpomene*[9,11]. Coloured blocks represent annotated gene regions on forward and reverse strands (see Supplementary Fig. 1 and Supplementary Table 3 for details). **b**, Association of SNP variation with mimicry polymorphism in a sample of 25 *silvana* and 34 *aurora*/*arcuella* individuals from a single population. Markers genotyped across rearrangements BP1 (pink) and BP2 (blue) show perfect association of SNP variation with wing pattern. No association was found in the flanking region from markers *Fox* (*Hn00106*) to *BmSuc* (*Hn00019*), or at 12 unlinked loci (green). The association decays more slowly in the direction of loci *Bm5536* (*GCP*), *Bm5586* (*NudC*) and *Bm5593* (*Srp68*). **c**, LD heat map. Perfect LD (genotypic correlation coefficient $r^2 = 1$) is found across 580 kb spanning the BP1 and BP2 rearrangements ($n = 59$). LD decays rapidly outside this interval, although strong within-marker LD remains at *HN00021*. Markers that are unlinked to *P* show little LD with each other or with *P*. **d**, Haplotype network for marker *LRR* (*Hn00024*) in the Yurimaguas population, coloured according to wing-pattern phenotype. Haplotype clades separated by seven fixed differences are in complete association with wing pattern, taking into account dominance relationships. Similar haplotype clades were found for all loci genotyped within *P*, and across the Amazon basin, but not for genes flanking *P* or in unlinked regions (Supplementary Figs 4 and 6).



©2011 Macmillan Publishers Limited. All rights reserved

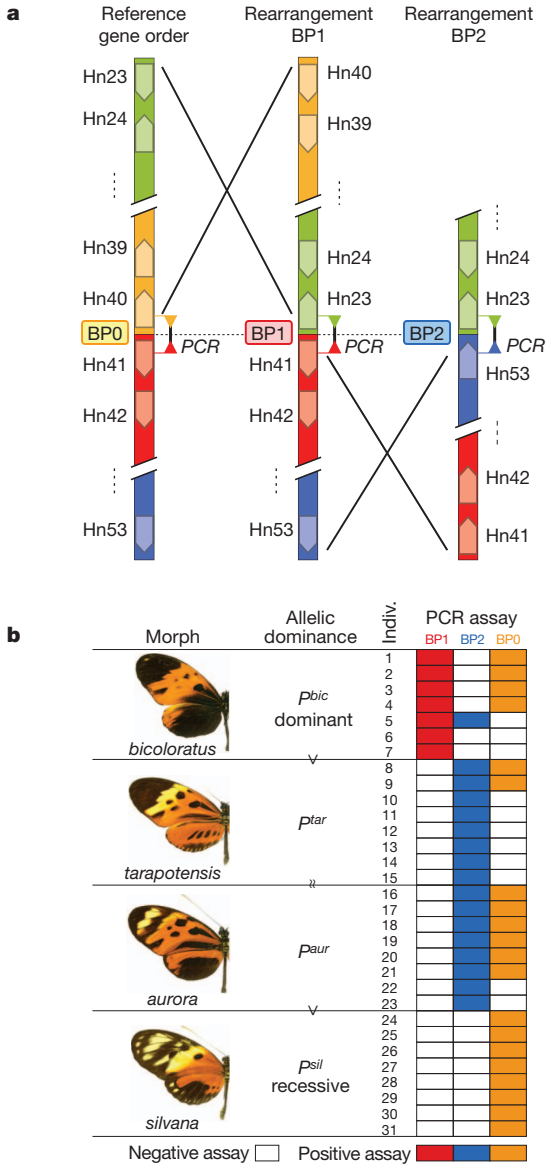



**Figure 3 | Chromosomal rearrangements associated with the supergene in natural populations.** **a**, Comparison of the gene orders found in the *H. numata* BAC library and wild populations. The rearrangements involve the 400-kb segment from genes *HM00023* to *HM00040* (*ERCC6*) (BP1, clones 24I10 and 45B17), and the adjacent 180-kb segment from gene *HN00041* (*penguin*) and *HN00053* (*lethal (2) giant larvae homologue*) (BP2, clone 38G4). Genes closest to the breakpoints are shown. **b**, Long-range PCR assays across alternative breakpoints (BP0, BP1 and BP2) in wild populations show a perfect association of the polymorphic gene orders with mimicry variation in four morphs from natural populations of Eastern Peru[15], following the dominance relationships (Fig. 1). The Yurimaguas population segregates primarily for *silvana* and *aurora*/*arcuella* forms, associated with BP0 and BP2, respectively. The Tarapoto population segregates mainly for *tarapotensis* and *bicoloratus* forms, associated with BP2 and BP1, respectively. This population also harbours recessive *illustris* alleles associated with BP0 (Supplementary Table 6).

gene rearrangements are evolutionarily derived and associated with the evolution of this locus in the *H. numata* lineage.

The results from BAC-clone sequencing were extended to natural populations by performing breakpoint-specific PCR on butterflies collected in the field (Fig. 3a and Supplementary Table 5). Long-range PCR analyses of 31 individuals of four morphs revealed a complete association between alternative rearrangements and specific wing-pattern phenotypes (Fig. 3b). Rearrangement BP1 was found in every *H. numata bicoloratus* and in no other morph, as expected for the dominant *P* allele, whereas BP2 was found in all intermediate dominant forms (*aurora* and *tarapotensis*). All recessive *silvana* individuals were BP0 homozygotes, confirming that the reference gene order segregates in *H. numata* populations. Some individuals with intermediate and dominant alleles had breakpoints that were lower in the dominance series, reflecting their heterozygosity. To confirm the association on a larger sample, short-range PCR assays primed closer to the breakpoints were performed on 156 individuals of six major morphs (Fig. 1 and Supplementary Tables 5–7). Short-range assays were highly consistent with long-range assays (Supplementary Table 6). Natural populations thus harbour at least three chromosomal arrangements in tight association with wing-pattern phenotypes. A short-range PCR product from BP1 was not amplified in two of the 32 *bicoloratus* individuals tested. This could reflect PCR failure due to sequence variation at BP1, or it may indicate that the BP1 breakpoint is not directly causative of the *bicoloratus* phenotype, despite being strongly associated with the causative variant(s).

To assess whether the recombination suppression that is observed in crosses operates at the population level, we estimated LD between nucleotide positions in and around *P*. A notable pattern of long-range LD was seen in tight association with the position of *P*, mapped from a survey of 17 markers sequenced in 59 individuals from the polymorphic population near Yurimaguas (Figs 1 and 2c). Complete LD was found between high-frequency single nucleotide polymorphisms (SNPs) across the 400-kb *P* interval, showing that long-range haplotypes are maintained across the supergene. This haplotype structure decays rapidly outside *P*: flanking markers lack the haplotype segregation found within *P* and show comparable levels of LD to the levels seen in 12 unlinked markers (Fig. 2c and Supplementary Fig. 3). Furthermore, within *P*, complete association was found between the divergent haplotypes and wing-pattern alleles. All recessive *silvana* individuals were homozygous for 39 SNPs in six genes across the *P* locus, whereas *aurora* individuals were heterozygous or homozygous for the alternative nucleotides (Fig. 2b, Supplementary Fig. 4 and Supplementary Table 5). Again, this complete association breaks down immediately outside *P* (Fig. 2b and Supplementary Figs 4 and 5). Therefore, the divergent haplotypes associated with wing-pattern alleles seem to be confined within the boundaries of the polymorphic rearrangements and the position of the supergene.

Finally, to confirm that the haplogroups are associated with phenotype and are not due to local population processes, we screened a population of *H. numata* from French Guiana, situated 2,900 km from Peru but harbouring phenotypically similar morphs (*silvana* and *numata*; Supplementary Table 5)[4]. The same two haplotype groups were found, and breakpoint PCR and haplotype clades were perfectly associated with corresponding phenotypes in both locations (Supplementary Fig. 6 and Supplementary Table 6). The association between mimicry polymorphism, local rearrangements and divergent haplotype blocks therefore seems to be conserved across the Amazonian range.

Our results show that the supergene is characterized by a 400-kb chromosomal block that is sharply structured into distinct haplotype clades separated by 1–4% divergence. These clades correspond to different wing-pattern alleles, segregate consistently with the allelic dominance and are associated with the chromosomal rearrangements (Fig. 2d and Supplementary Figs 3 and 4). This situation stands in stark contrast to the related species *H. melpomene* and *H. erato*, in which several independent wing-patterning loci are under directional selection, and in which markers surveyed in the orthologous region showed no fixed nucleotide differences between colour-pattern races, and no long-range LD[10,24]. The *P* supergene architecture, associated with mimicry polymorphism under balancing selection in *H. numata*[4,15], is thus evolutionarily characterized by non-recombining co-adapted blocks that capture distinct wing-pattern genes that are known to recombine in other species of the clade[10,24].

In summary, the four strands of evidence, namely fine-scale comparative mapping, association of three chromosomal arrangements with morphs, SNP–phenotype association at *P*, and long-range LD



©2011 Macmillan Publishers Limited. All rights reserved

and divergent haplotype clades, all indicate that the rearrangements lock together distinct elements involved in wing-pattern evolution, providing long-awaited evidence for a situation that may apply to the evolution of supergenes in other systems[7,16–19]. Together, the data begin to explain how distinct loci that control geographical variation in some species[10,11] can become locked together in others. Our results highlight the role that structural variation can have in generating co-adapted gene complexes involved in adaptation and speciation[8,12,19,20,25], and open the way to the study of their functional integration.

The next challenge will be to identify the sites that are causally involved in the elements of pattern variation in *H. numata*, a goal hindered by reduced recombination within *P*, which limits the power of fine mapping to dissect the locus. Instead, functional studies of genes in the interval and analysis of rare recombinants will be important. Notably, *P* seems to be a hotspot of adaptation in several other species[26], including the melanic peppered moth[27], and, along with other colour-pattern regions, is linked to assortative mating and speciation in other *Heliconius* species[28–30]. Unravelling the genetic nature of the elements that contribute to the supergene will therefore be key to understanding the mechanisms and sequence of events underlying the clustering of adaptive traits in rearrangements that are associated with ecological divergence.

## METHODS SUMMARY

A BAC library was screened with radiolabelled probes derived from BAC end-sequences during chromosome walking. Clone sequences, obtained by capillary sequencing, were annotated with the aid of a Roche 454-sequenced cDNA library from the developing wing[11]. Genomic sequences were aligned to a *H. melpomene* reference[11] to detect chromosomal rearrangements. Crossover events were fine-mapped using broods described elsewhere[9]. Recombination estimates from other genomic regions were derived by scoring crossing over along two unlinked clones, using markers separated by known physical distance.

PCR-based diagnostics for the rearrangement breakpoints BP1 and BP2 were designed from *H. numata* BAC sequences. BP0 was designed from the reference *H. melpomene* sequence. Long-range PCRs were primed in exons flanking each breakpoint, and verified by sequencing. Short-range PCRs were primed closer to the breakpoints in unique intergenic regions (Supplementary Table 7).

Genetic markers were sequenced by capillary sequencing. LD was estimated from unphased sequences using genotypic correlation tests between all pairs of variable sites. The association between wing-pattern and SNPs was tested using Fisher's exact tests, taking into account the dominance hierarchy of wing-pattern alleles. Haplotype networks were constructed by parsimony using Network.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 14 March; accepted 28 June 2011.**
**Published online 14 August 2011.**

1. Mather, K. The genetical architecture of heterostyly in *Primula sinensis*. *Evolution* **4,** 340–352 (1950).
2. Garber, R. J. & Quisenberry, K. S. The inheritance of length of style in buckwheat. *J. Agric. Res.* **34,** 181–183 (1927).
3. Clarke, C. A., Sheppard, P. M. & Thornton, I. W. B. The genetics of the mimetic butterfly *Papilio memnon*. *Philos. Trans. R. Soc. Lond. B* **254,** 37–89 (1968).
4. Brown, K. S. & Benson, W. W. Adaptive polymorphism associated with multiple Müllerian mimicry in *Heliconius numata*. *Biotropica* **6,** 205–228 (1974).
5. Nijhout, H. F. Polymorphic mimicry in *Papilio dardanus*: mosaic dominance, big effects, and origins. *Evol. Dev.* **5,** 579–592 (2003).
6. Murray, J. & Clarke, B. Supergenes in polymorphic land snails—examples from genus *Partula*. *Genetics* **74,** S188–S189 (1973).
7. Kirkpatrick, M. & Barton, N. Chromosome inversions, local adaptation and speciation. *Genetics* **173,** 419–434 (2006).
8. Manfield, I. W. *et al.* Molecular characterization of DNA sequences from the *Primula vulgaris S*-locus. *J. Exp. Bot.* **56,** 1177–1188 (2005).
9. Joron, M. *et al.* A conserved supergene locus controls colour pattern diversity in *Heliconius* butterflies. *PLoS Biol.* **4,** e303 (2006).
10. Baxter, S. W. *et al.* Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in the *Heliconius melpomene* clade. *PLoS Genet.* **6,** e1000794 (2010).
11. Ferguson, L. *et al.* Characterization of a hotspot for mimicry: Assembly of a butterfly wing transcriptome to genomic sequence at the *HmYb/Sb* locus. *Mol. Ecol.* **19,** 240–254 (2010).
12. Hoffmann, A. A. & Rieseberg, L. H. Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annu. Rev. Ecol. Evol. Syst.* **39,** 21–42 (2008).
13. Pinho, C. & Hey, J. Divergence with gene flow: Models and data. *Annu. Rev. Ecol. Evol. Syst.* **41,** 215–230 (2010).
14. Clark, R. *et al.* Colour pattern specification in the Mocker swallowtail *Papilio dardanus*: the transcription factor invected is a candidate for the mimicry locus H. *Proc. R. Soc. B* **275,** 1181–1188 (2008).
15. Joron, M., Wynne, I. R., Lamas, G. & Mallet, J. Variable selection and the coexistence of multiple mimetic forms of the butterfly *Heliconius numata*. *Evol. Ecol.* **13,** 721–754 (1999).
16. Turner, J. R. G. in *The biology of butterflies* Vol. 11 (eds Vane-Wright, R. I. & Ackery, P. R.) 141–161 (Academic, 1984).
17. Charlesworth, D. & Charlesworth, B. Theoretical genetics of Batesian mimicry. II. Evolution of supergenes. *J. Theor. Biol.* **55,** 305–324 (1975).
18. Alvarez, G. & Zapata, C. Conditions for protected inversion polymorphism under supergene selection. *Genetics* **146,** 717–722 (1997).
19. Hatadani, L. M., Baptista, J. C. R., Souza, W. N. & Klaczko, L. B. Colour polymorphism in *Drosophila mediopunctata*: genetic (chromosomal) analysis and nonrandom association with chromosome inversions. *Heredity* **93,** 525–534 (2004).
20. Lowry, D. B. & Willis, J. H. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* **8,** e1000500 (2010).
21. Jones, R. T., Salazar, P., ffrench-Constant, R. H., Jiggins, C. D. & Joron, M. Evolution of a mimicry supergene from a multilocus architecture. *Proc. R. Soc. B*. doi:10.1098/rspb.2011.0882 (2011).
22. Papa, R. *et al.* Highly conserved gene order and numerous novel repetitive elements in genomic regions linked to wing pattern variation in *Heliconius* butterflies. *BMC Genomics* **9,** 345 (2008).
23. Pringle, E. G. *et al.* Synteny and chromosome evolution in the lepidoptera: Evidence from mapping in *Heliconius melpomene*. *Genetics* **177,** 417–426 (2007).
24. Counterman, B. A. *et al.* Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in *Heliconius erato*. *PLoS Genet.* **6,** e1000796 (2010).
25. Dobzhansky, T. Genetics of natural populations. XIV. A response of certain gene arrangements in the third chromosome of *Drosophila pseudoobscura* to natural selection. *Genetics* **32,** 142–160 (1947).
26. Saenko, S. V., Brakefield, P. M. & Beldade, P. Single locus affects embryonic segment polarity and multiple aspects of an adult evolutionary novelty. *BMC Biol.* **8,** 111 (2010).
27. van't Hof, A. E., Edmonds, N., Dalíková, M., Marec, F. & Saccheri, I. J. Industrial melanism in British peppered moths has a singular and recent mutational origin. *Science* **332,** 958–960 (2011).
28. Kronforst, M. R. *et al.* Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. *Proc. Natl Acad. Sci. USA* **103,** 6575–6580 (2006).
29. Chamberlain, N. L., Hill, R. I., Kapan, D. D., Gilbert, L. E. & Kronforst, M. R. Polymorphic butterfly reveals the missing link in ecological speciation. *Science* **326,** 847–850 (2009).
30. Merrill, R. M., Van Schooten, B., Scott, J. A. & Jiggins, C. D. Pervasive genetic associations between traits causing reproductive isolation in *Heliconius* butterflies. *Proc. R. Soc. B* **278,** 511–518 (2010).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank M. Blaxter and D. Charlesworth for advice throughout the study; The GenePool and S. Humphray for DNA sequencing; S. Kumar and A. Papanicolaou for bioinformatics support; M. Beltrán, A. Bulski, M. Veuille and the Botanique-Entomologie-Mycologie molecular facility (BoEM) for laboratory support; S. Johnston for genome-size estimates in *H. numata*; D. Obbard for providing R scripts; M. Abanto, S. Gallusser, C. Ramírez, L. de Silva, J. Barbut, B. Gilles and G. Lamas for help with butterfly rearing, fieldwork and collecting permits; and the Peruvian National Institute of Natural Resources (INRENA) for granting collecting and export permits (076-2007-INRENA-IFFS-DCB). Fieldwork in French Guiana was supported by a CNRS 'Nouragues Research Grant'. This work was supported by an EMBO long-term fellowship (ALTF-431-2004), EMBO-matching funds from NWO (Netherlands), a Royal Society University Research Fellowship (516002.K5917/ROG), a CNRS grant (ATIP Biodiversité 2008, France) and a European Research Council Starting Grant (ERC-Stg 'MimEvol') to M.J., a BBSRC grant (BBE0118451) to C.D.J. and R.H.ff.-C., a Leverhulme Trust grant (F/00144AY) to R.H.ff.-C., and a Royal Society University Research Fellowship and a Leverhulme Research Leadership grant to C.D.J.

**Author Contributions** M.J., C.D.J. and R.H.ff.-C. designed the study and contributed to all stages of the project. M.J., L. Frezal and R.T.J. performed the principal experiments and data analysis, with assistance from N.L.C., S.W.B., S.F.L., M.B., C.S., L. Ferguson, C.R.H., A.W. and P.A.W. BAC clone sequencing was carried out by C.D., R.G., C.L., R.C., H.B., S.S., J.R., M.C.J. and M.A.Q. M.J., A.W., C.D.J. and R.H.ff.-C. co-wrote the manuscript with input from all authors.

**Author Information** GenBank accessions for BAC clone sequences: FP885863, FP476061, FP565803, FP476023, CU856181, FP885878, FP476047, FP885857, CU856182, CU655868, FP885879, FP885861, FP885880, FP885855, CU914733, FP475989, CU655869, CU914734, CU633161, CU638865, CU856175, FP884220 and FP236755. Accessions for 1364 marker sequences: JN173798–JN175161. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.J. (joron@mnhn.fr).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Chromosome walking.** A BAC library was constructed by Amplicon Express from five individual larvae from a polymorphic population of *H. numata* segregating for the forms *bicoloratus*, *tarapotensis* and *arcuella*. The BAC library has an estimated 119-kb average insert size and ×7 coverage of the 319-Mb genome. It was printed onto nylon filters and screened using 11 radiolabelled PCR probes designed to span the orthologous genomic sequences from an *H. melpomene* BAC library[9,11]. The BAC library was also fingerprinted by restriction digest, and overlapping clones were predicted from analysis of a fingerprinted contig (FPC) database. Fifty-three BAC clones were identified and tested for positive amplification of the original probes. Nineteen *H. numata* clones were sequenced by the Wellcome Trust Sanger Institute to high-throughput-genomic (HTG) phase 3 quality, totalling 2.9 Mb of overlapping genomic sequence. Because of the multiple chromosomal arrangements found in this region, and the fact that the clones sequenced come from chromosomes with differing gene orders as a consequence, it is impossible to reconstruct a single linear tilepath of clones at this stage. On the basis of reciprocal BLASTs of the clone sequences, the tilepath covers 0.95 Mb centred on *P*, consisting of four 'floating' contigs, two of which lie within the recombination interval of *P* and show different gene orders. A gap in the BAC chromosomal walk, estimated to be approximately 150 kb, remains between clones 7C9 and 14K13 despite extensive screening of the library. However, the gene markers *ARP-like* (*Hm00028*) and *Pros54* (*Hm00030*), predicted from the reference *Heliconius* sequence to lie in the middle of this gap, do indeed map in full linkage with other markers on either side of the gap, indicating the gene content of the gap is probably conserved. Two additional contigs of clones were identified and sequenced approximately 350 kb and 900 kb from the end of the tilepath further down the chromosome (Fig. 2a and Supplementary Table 2). Sequence similarity was plotted using the software Vista[31,32], with a 70% identity threshold and a 100-bp sliding window (Supplementary Fig. 2).

**Linkage mapping.** Male-informative markers were used to score crossing over events between markers linked and unlinked to *P*. Female Lepidoptera have achiasmatic meiosis, so crossing over only occurs during gametogenesis in the male parent. Thirty-six PCR-based markers were designed from BAC sequences and genotyped by visualization of differences in amplicon size, by restriction-fragment length polymorphism, or by sequence variation, in 366 individuals from six mapping families (B377, B465, B472, B502, B523 and A298 (refs 9, 11)). This was used to circumscribe the *P* supergene and orientate the tilepath by recombination. Special emphasis was given to the regions near the first crossover on either side of the supergene, for a precise positioning of the mapping boundary. The three recombinants with *P* on one side, and two recombinants on the other side, ensured that *P* was circumscribed and in full linkage with markers in an interval estimated to be about 400 kb (using a 0.92 scaling factor for *H. numata* versus *H. melpomene* genomic sequence, on the basis of sequences available for this region). The same procedure was used to score recombination events in unlinked regions using two sequenced BAC clones (bHN20L19 and bHM7E22) from unlinked chromosomes and three mapping families (Supplementary Table 1).

**BAC gene annotation.** Transcriptome sequencing, assembly and genomic annotation were performed as described elsewhere[11]. Briefly, annotation and gene predictions in all sequenced BACs (Supplementary Fig. 1 and Supplementary Table 3) were carried out with the Maker annotation pipeline[33], using 600K expressed sequenced tags generated with Roche 454 FLX pyrosequencing (200-base average read length) and assembled using the Mira assembler[34] to yield 24,992 objects. Complementary DNA was generated from wing-disc tissue extracted from juvenile stages from the Peruvian populations used throughout this study. Large variations in non-coding content, as well as in the presence and identity of transposable elements, were noted between clones within *H. numata* (Supplementary Table 4), causing some *H. numata* clones (for example, 14K13) to show a higher overall similarity to the orthologous *H. melpomene* sequences (bHM29B7) than to other orthologous *H. numata* clones (46M23 and 38G4). This indicates that the two *H. numata* chromosomal segments with differing gene orders may be anciently derived.

**Sequencing and breakpoint analysis.** Annotation of the BACs was used to identify exons within and outside the *P* mapping interval; exon markers were PCR-amplified and direct-sequenced for population-genetic analysis (Supplementary Table 8). Unlinked markers were chosen following the method in ref. 35. A total of 17 amplicons lying on linkage group 15, both within and outside the region containing *P*, plus 13 amplicons on unlinked chromosomes, were sequenced in 48–144 individuals (Supplementary Table 5) from a polymorphic population in eastern Peru. This populations segregates predominantly for the forms *silvana* and *aurora* (including sub-variants *isabellinus* and *elegans*). In subsequent analyses, we combined the *aurora* variants under the single class *aurora*, because the numerous phenotypic gradations that are found between these forms contrast with the clearly distinct phenotypic form *silvana*[4,15,36]. Furthermore, PCR assays of the rearrangement breakpoints and the absence of diagnostic nucleotide differences also indicated that these variants could be combined.

Breakpoint PCR assays were carried out with primers designed from exonic sequences on either side of each breakpoint (Supplementary Table 7). Fragments were amplified from 31 individuals with four different wing patterns (*bicoloratus*, *tarapotensis*, *aurora* and *silvana*) using long-range PCR (Qiagen) following the manufacturer's conditions, and were end-sequenced to confirm fragment identity. A second, larger sample of 201 individuals (including 161 from Peru and 40 from French Guiana) was assayed by standard (short-range) PCR amplification (Fermentas DreamTaq) using primers positioned closer to the breakpoints in unique non-coding DNA (Supplementary Table 7). Capillary sequencing of markers and breakpoint assays was performed on an ABI 3730 capillary sequencer with BigDye chemistry by the University of Edinburgh sequencing service (http://genepool.bio.ed.ac.uk/). Sequencing ambiguities were resolved manually using CodonCode Aligner (http://www.codoncode.com).

**Phenotype-by-genotype associations.** Association between genotype and phenotype was estimated for all polymorphic sites (SNPs) within each marker, taking into account the dominance relationships between different morphs. The Amazonian form of *H. numata* termed *silvana* (widespread from French Guiana to the Andean foothills) and its geographic replacement *illustris* (Andean valleys) are recessive to all other forms[4,9,36], which predicts that all *silvana* and *illustris* individuals are homozygous at *P*, whereas individuals of other, co-occurring forms can be either heterozygous or homozygous for a different allele. We estimated the association between genotype and phenotype by testing the hypothesis that the major allele should be homozygous in *silvana* individuals and heterozygous or absent in *aurora* individuals. For each polymorphic site with a minor allele frequency of at least 0.1, we calculated the proportion of individuals whose genotype conformed to this hypothesis and tested the null hypothesis of no phenotype–genotype association using Fisher's exact test. This method allows testing of the phenotype-by-genotype association while using knowledge of the dominance relationships between *P* alleles.

**Linkage disequilibrium (LD) across the *P* supergene.** LD was assessed for all pairs of polymorphic sites within and surrounding the *P* region, as well as in 12 unlinked markers (Fig. 2c) in 59 specimens from a polymorphic population near Yurimaguas, eastern Peru (Supplementary Table 5). We used genotypic correlation-based testing for LD with several alleles with unknown phase between sites[37]. Significance was tested by a permutation test for the genotypic correlation statistic $r^2$ (ref. 37).

**Haplotype networks.** The phase of SNP variation was determined by analysing the segregation of alleles from parent to offspring in markers sequenced from the mapping families, which originate from the same populations as the population samples. Haplotypes from the population samples were inferred by coalescent-based Bayesian methods using the PHASE 2.1 algorithm[38], optimized by including sequences with known phase from mapping families. The high level of LD across the region ensured a robust inference of haplotypic diversity for all markers. Haplotype networks were constructed by parsimony using the Network package[39,40] (http://www.fluxus-engineering.com/). The level of genetic differentiation between the *silvana* and non-*silvana* groups was estimated in DNAsp[41] using the Fst statistics[42], and its significance was tested by permutation (Supplementary Fig. 5). Fst was not used to test for genetic differentiation among different populations, but rather to assess genetic differentiation between two groups of individuals in a single population (*silvana* versus non-*silvana*). Even high levels of genetic differentiation do not indicate the absence of random mating, but rather that these sites are linked and/or that they co-vary with the functional loci that determine the different mimetic morphs.


31. Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M. & Dubchak, I. VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* **32,** W273–W279 (2004).
32. Mayor, C. *et al.* VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* **16,** 1046 (2000).
33. Cantarel, B. L. *et al.* MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18,** 188–196 (2008).
34. Chevreux, B. *et al.* Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res.* **14,** 1147–1159 (2004).
35. Salazar, C., Jiggins, C. D., Taylor, J. E., Kronforst, M. R. & Linares, M. Gene flow and the genealogical history of *Heliconius heurippa*. *BMC Evol. Biol.* **8,** 132 (2008).
36. Brown, K. S. An illustrated key to the silvaniform *Heliconius* (Lepidoptera: Nymphalidae) with descriptions of new subspecies. *Trans. Am. Entomol. Soc.* **102,** 373–484 (1976).
37. Zaykin, D. V., Pudovkin, A. & Weir, B. S. Correlation-based inference for linkage disequilibrium with multiple alleles. *Genetics* **180,** 533–545 (2008).
38. Stephens, M. & Donnelly, P. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am. J. Hum. Genet.* **73,** 1162–1169 (2003).


©2011 Macmillan Publishers Limited. All rights reserved

39. Polzin, T. & Daneshmand, S. V. On Steiner trees and minimum spanning trees in hypergraphs. *Oper. Res. Lett.* **31,** 12–20 (2003).
40. Forster, P., Torroni, A., Renfrew, C. & Rohl, A. Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. *Mol. Biol. Evol.* **18,** 1864–1881 (2001).
41. Librado, P. & Rozas, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25,** 1451–1452 (2009).
42. Hudson, R. R., Slatkin, M. & Maddison, W. P. Estimation of levels of gene flow from DNA-sequence data. *Genetics* **132,** 583–589 (1992).

©2011 **Macmillan Publishers Limited. All rights reserved**

LETTER

doi:10.1038/nature10347

# Cell-to-cell spread of HIV permits ongoing replication despite antiretroviral therapy

Alex Sigal[1], Jocelyn T. Kim[1,2], Alejandro B. Balazs[1], Erez Dekel[3], Avi Mayo[3], Ron Milo[4] & David Baltimore[1]

**Latency and ongoing replication[1] have both been proposed to explain the drug-insensitive human immunodeficiency virus (HIV) reservoir maintained during antiretroviral therapy. Here we explore a novel mechanism for ongoing HIV replication in the face of antiretroviral drugs. We propose a model whereby multiple infections[2,3] per cell lead to reduced sensitivity to drugs without requiring drug-resistant mutations, and experimentally validate the model using multiple infections per cell by cell-free HIV in the presence of the drug tenofovir. We then examine the drug sensitivity of cell-to-cell spread of HIV[4–7], a mode of HIV transmission that can lead to multiple infection events per target cell[8–10]. Infections originating from cell-free virus decrease strongly in the presence of antiretrovirals tenofovir and efavirenz whereas infections involving cell-to-cell spread are markedly less sensitive to the drugs. The reduction in sensitivity is sufficient to keep multiple rounds of infection from terminating in the presence of drugs. We examine replication from cell-to-cell spread in the presence of clinical drug concentrations using a stochastic infection model and find that replication is intermittent, without substantial accumulation of mutations. If cell-to-cell spread has the same properties *in vivo*, it may have adverse consequences for the immune system[11–13], lead to therapy failure in individuals with risk factors[14], and potentially contribute to viral persistence and hence be a barrier to curing HIV infection.**

Current antiretroviral therapy (ART) does not cure HIV infection because low-level viraemia persists from virus reservoirs that are insensitive to ART[1]. The reservoirs may be long-lived infected cells, cells with latent virus, ongoing cycles of infection termed ongoing replication, or a combination of sources[1]. How ongoing replication might take place in the face of ART has remained unclear. If ART succeeds in decreasing ongoing HIV replication to very low levels, why does it not eliminate replication completely? Here we explore a novel mechanism for ongoing HIV replication in the presence of ART.

Multiple infections of one cell may propagate at drug concentrations where infection by single particles would die out: if more virions are transmitted per cell, the probability that at least one of the virions escapes the drug should increase (Fig. 1a). To model the effect of multiple infections on drug sensitivity (Supplementary Theory, section 1), we assume infections by individual virions are independent events, each with a probability of escaping the drug and succeeding in infecting the cell. To quantify infection sensitivity to drugs, we introduce the transmission index ($T_X$), which we define as the fraction of cells infected in the presence of drug ($I_d$) divided by the fraction of cells infected in the absence of drug ($I$). Given: (1) a multiplicity of infection of $m$ infectious units per cell, where $m$ is defined as the product of virus particle number and the probability of infection per virus particle; (2) a concentration of antiretroviral agent $d$ that reduces $m$ by factor $f(d)$, where $f(d) \geq 1$. Under these conditions, the transmission index is:

$$T_X = \frac{I_d}{I} = \frac{1-e^{-m/f(d)}}{1-e^{-m}} \quad (1)$$

$T_X$ has two important limiting regimes: $m \ll 1$, in which case $T_X \approx 1/f(d)$ and $m/f(d) \gg 1$, in which case $T_X \approx 1$. In the first case, where few viruses infect each cell, the infection is sensitive to the effect of the drug, whereas in the second, where many viruses infect each cell, the infection is insensitive.

To test this, we infected the highly infection-permissive MT-4 T-cell line with cell-free HIV encoding yellow fluorescence protein (YFP)[15] at low (0.2) and high (100) $m$ in the presence of tenofovir (TFV), a nucleotide reverse transcriptase inhibitor. We determined infected cell number by YFP fluorescence (Supplementary Fig. 1) and observed that infection with cell-free virus at low $m$ was sensitive to TFV across the range of concentrations used. At high $m$, infection was insensitive to low and intermediate TFV concentrations (Fig. 1b), supporting the model. Thus, multiple cell-free HIV infections per cell recapitulate the insensitivity to drug of an HIV reservoir.

Multiple infections occur *in vivo*[2,16] and in culture[8,10] and are thought to be associated with cell-to-cell spread[2,8–10], a directed transmission mode that minimizes the number of virus particles failing to

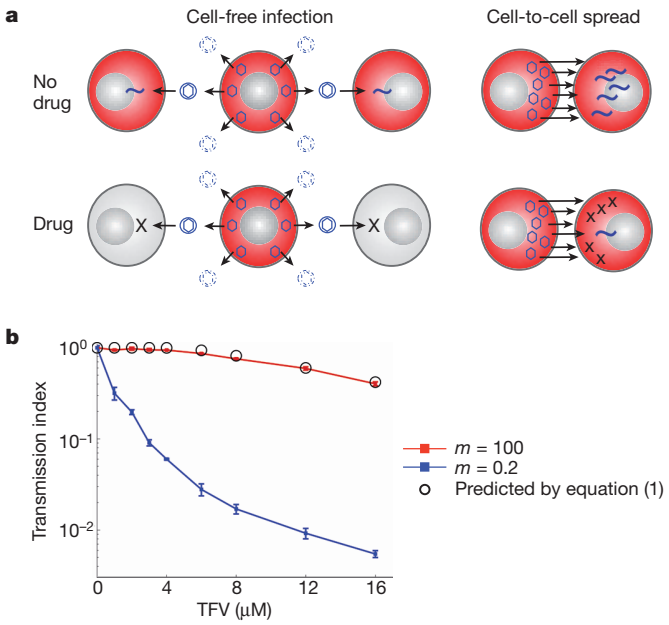



**Figure 1 | Multiple infections per cell decrease sensitivity to drug.** **a**, Hypothesis. Red circles indicate infected cells, arrows indicate transmissions, hexagons or hexagons surrounded by circles indicate viruses, broken circles indicate degraded viruses, crosses indicate viruses blocked by drug and wavelets indicate successful infection. **b**, MT-4 cells were pre-incubated with TFV and infected with HIV coding for YFP. Infection multiplicity $m$ was 0.2 (blue squares) or 100 (red squares). Lines are a guide for the eye. Mean ± standard deviation (s.d.) of replicates ($n = 3$). Circles represent calculated values of $T_X$ at $m = 100$ according to equation (1) with $f(d)$ at each drug concentration determined empirically at $m = 0.2$.

[1]Division of Biology, California Institute of Technology, Pasadena, California 91125, USA. [2]Division of Infectious Diseases, Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, California 90095, USA. [3]Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 76100, Israel. [4]Department of Plant Sciences, Weizmann Institute of Science, Rehovot 76100, Israel.



©2011 Macmillan Publishers Limited. All rights reserved

reach the target cell. We therefore used co-culture with infected cells to generate cell-to-cell spread and compared drug sensitivity to infection with cell-free virus. Infection by co-culture occurs both by cell-free virus shed by infected donor cells and by cell-to-cell spread. Administration of cell-free virus lacks a cell-to-cell component—the measured average virus cycle time (1.7 days; Supplementary Fig. 2) would rarely permit cell-free virus infected cells to complete a second round of infection during the experiment (2 days). Therefore, we compared cell-free virus infection and the combination of cell-free virus infection and cell-to-cell spread resulting from co-culture. We used drugs that act far downstream of entry, to ensure any differences between cell-to-cell and cell-free infection are not due to factors that physically inhibit drug action in cell-to-cell spread.

We infected peripheral blood mononuclear cells (PBMCs) in the presence or absence of TFV by co-culture or using cell-free virus. To separate donor from target cells in co-culture, we used HLA-A2-negative donor cells and HLA-A2-positive targets (Supplementary Fig. 3a). Two days post-infection, we determined the fraction of target cells infected using p24 intracellular staining of HLA-A2-positive PBMCs (Fig. 2a, top panel, controls in Supplementary Fig. 3b). Co-culture dramatically decreased sensitivity to drug: TFV decreased cell-free infection ~30-fold but caused less than a twofold decrease of co-culture infection (Fig. 2b). The decline in HLA-A2 expression in the target cells after infection (Supplementary Fig. 3b) is consistent with observations that productive HIV infection downregulates HLA[17].

We also used Rev-CEM[18] reporter T cells as targets. These cells express green fluorescent protein (GFP) in the presence of HIV early proteins Tat and Rev (Supplementary Fig. 4). To infect Rev-CEM cells, we used either cell-free HIV or co-culture with infected MT-4 cells engineered to be >99% mCherry positive (Supplementary Fig. 5). We excluded GFP/mCherry double-positive cells from the analysis to avoid scoring fused cells as infected (Supplementary Figs 6 and 7). This underestimates co-culture infection because it excludes unfused cell doublets in the process of virus exchange.

To block infection, we applied TFV and the non-nucleoside reverse transcriptase inhibitor efavirenz (EFV) (Fig. 2a, bottom panel, Supplementary Fig. 7). At the highest concentrations used, co-culture $T_X$ was over sixfold higher than cell-free infection $T_X$ (Fig. 2b). The trend was similar when donors were PBMCs or Rev-CEM cells (Supplementary Fig. 8). Co-culture $T_X$ was lower than in PBMC-to-PBMC transmission, suggesting that target cells have an important role in cell-to-cell spread efficiency. The lower drug sensitivity in co-culture was not due to secreted donor cell factors that decrease the susceptibility of target cells to drugs (Supplementary Fig. 9).

We next determined the number of infectious units ($m$) transmitted. For co-culture, $m$ was previously proposed to have a two-peaked Poisson distribution, one peak corresponding to cell-free virus or some low virus cell-to-cell transmissions, and the second to high virus number transmissions[3,9]. We fit a two-peaked Poisson and other distributions to the data (Supplementary Theory, section 2). The two-peaked Poisson fit the data best (Fig. 2b, dotted line, Supplementary Fig. 10). The first peak mean was ~1 infectious unit for both drugs, with 94% and 97% of infections in this peak for TFV and EFV, respectively. The second peak mean was 73 (TFV) and 175 (EFV), with the remaining 6% and 3% of infections in this peak. This predicts that whereas most infections are cell-free or low virus cell-to-cell transmissions, a minority involve very large numbers of viruses. This might seem to imply large numbers of integrations in the absence of drug in the high virus number subset. Arguing against this is our observation of a significantly increased cell death rate with increasing numbers of multiple infections in the absence of drugs (data not shown). Inter-virus interference, such as downregulation of CD4 receptors[19], may also limit provirus number.

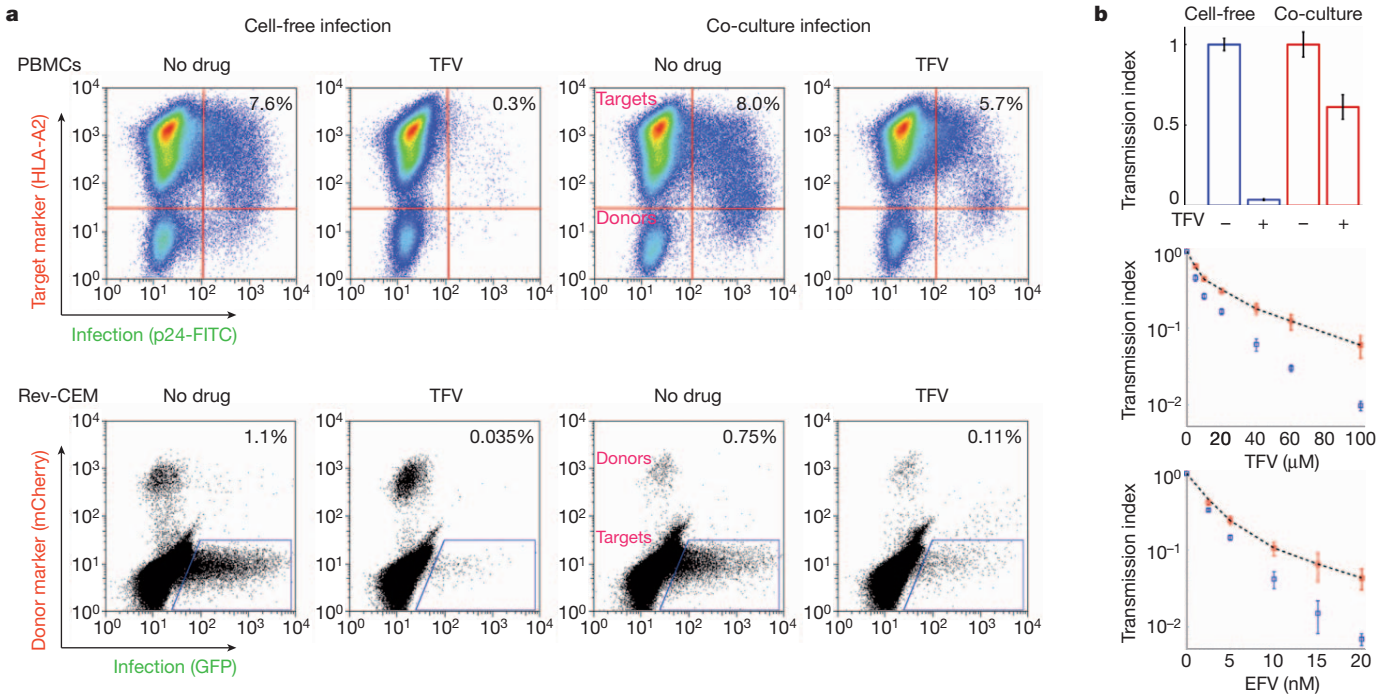



**Figure 2 | Cell-to-cell spread reduces sensitivity to drugs.** **a**, Top, infection of HLA-A2-positive PBMC targets with cell-free virus (left two plots) or infected HLA-A2-negative PBMC donors (right two plots) in the absence or presence of 10 μM TFV. $x$-axis is p24, $y$-axis HLA-A2 status. Bottom, the number of GFP-positive Rev-CEM cells after infection with cell-free virus (left two plots) or infected MT-4mCherry donors (right two plots) in the absence or presence of 60 μM TFV. $x$-axis is GFP, $y$-axis is mCherry fluorescence. **b**, Transmission index when infection source was cell-free HIV (blue bars or squares) or co-culture with HIV-infected donor cells (red bars or squares). Mean ± s.d. ($n = 3$). Top graph is PBMCs with TFV, middle graph is Rev-CEM cells with TFV, bottom graph is Rev-CEM cells with EFV. Black dashed line is best fit of $m$ with a two-peaked Poisson distribution described by $p(m;a,\mu_1,\mu_2) = (1-a)e^{-\mu_1}\mu_1^m/m! + ae^{-\mu_2}\mu_2^m/m!$, where $\mu_1$ and $\mu_2$ are the means of the first and second peak respectively, and $a$ is the fraction of transmissions that fall within the second peak. For TFV, $\mu_1 = 1.1$, $\mu_2 = 73$, $a = 0.06$. For EFV, $\mu_1 = 0.8$, $\mu_2 = 175$, with $a = 0.03$. Root mean squared error was 0.01 (EFV) to 0.02 (TFV).



©2011 Macmillan Publishers Limited. All rights reserved

To investigate whether cell-to-cell spread can lead to HIV replication through multiple virus cycles with ART, we measured the replication ratio ($R$), defined as fold change in the number of infected cells per virus cycle under conditions where target cells are not limiting: $(I_k/I_0)^{1/k}$. Here $k$ is the number of elapsed virus cycles, $I_k$ is the number of infected cells at virus cycle $k$, and $I_0$ is the number of infected cells at the start. For expanding infections $R > 1$, whereas infections with $R < 1$ ultimately terminate[20,21]. Although this assumes synchronized virus cycles, we simulated desynchronization and observed that its effect was negligible at the measured variability in cycle lengths (Supplementary Fig. 11).

To measure $R$, we tracked infection daily (Methods) in the absence of drug, with 100 μM TFV, or with a combination of EFV, TFV and the nucleoside reverse transcriptase inhibitor emtricitabine (FTC) at their clinical maximum plasma concentrations ($C_{max}$: 10 μM EFV, 2 μM TFV and 10 μM FTC[22]). The fraction of infected cells was kept low to ensure that target cells were not limiting. $R_0$, $R_{TFV}$ and $R_{C_{max}}$, the replication ratios with no drug, TFV or at $C_{max}$, were fitted from the data (Fig. 3a, dashed lines). They were 65, 2.5 and 0.95, respectively. $R_{TFV}$ was significantly greater than 1 ($P < 0.01$), indicating an expanding infection. $R_{C_{max}}$ was slightly lower than 1 in all experiments (Supplementary Fig. 12), indicating an infection slightly below the expansion threshold.

We compared experimentally obtained replication ratios with those predicted for the same drug concentrations if cell-free infection were the only infection route (Supplementary Theory section 3 and Supplementary Fig. 13). We obtained $R_{TFV} = 1.1$ and $R_{C_{max}} = 0.60$ values in this case (Fig. 3a). The predicted $R$ with no replication, resulting solely from infected cell half-life, was 0.46 (Fig. 3a). Predicted cell-free replication ratios were significantly lower ($P < 0.02$ for TFV, $P < 0.01$ for $C_{max}$) than ratios experimentally obtained from co-culture.

Given the lack of evolution in the plasma in individuals with HIV successfully suppressed by drugs[23,24], ongoing replication can occur if: (1) it is compartmentalized to other locations[25,26], (2) if it is intermittent; (3) the circulating virus is at a fitness maximum[24]; or some combination of these factors. We obtained $R_{C_{max}} = 0.95$. If this is extrapolated *in vivo*, it follows that ongoing replication cannot persist independently but may have a role if it interacts with another reservoir that primes replication[27]. To examine this scenario, we performed a stochastic simulation (Methods). As expected for intermittent replication, every infection chain that starts from the introduction of an infected cell from a different reservoir—for example, reactivation from latency—terminates (Supplementary Fig. 14). A constant input of one infected cell per virus cycle results in a steady state where substantial numbers of newly infected cells are generated, but the average number of mutations anywhere on the HIV genome per infected cell is low (~1; Fig. 3b). Because each infection chain is independent, these mutations are expected to be sporadic and not linked by temporal structure.

Evidence for ongoing replication during ART derives from the decrease in virus decline rates[28], some HIV sequence divergence[29] and long terminal repeat circle formation when the integrase inhibitor raltegravir is included in drug regimens[11]. At least in some individuals, antiretroviral suppression is close to the ongoing replication threshold: a mutation conferring very low-level resistance to EFV at therapy initiation[30] is sufficient to cause ongoing replication, as indicated by increased virological failure risk[14]. Our data indicate that cell-to-cell spread is a likely source of intermittent ongoing replication in the face of ART, and that this is a consequence of some cell-to-cell infections transmitting virus numbers much in excess of what is required to infect a cell in the absence of ART. The large transmitted dose strongly decreases the probability that every transmitted virus will be inhibited by the drugs, and therefore greatly weakens their effect. This replication may adversely affect the immune system, increasing activation[11,12] and cell death[13], and could potentially contribute to the maintenance of an HIV reservoir in locations such as lymphoid tissue where cell-to-cell spread occurs.

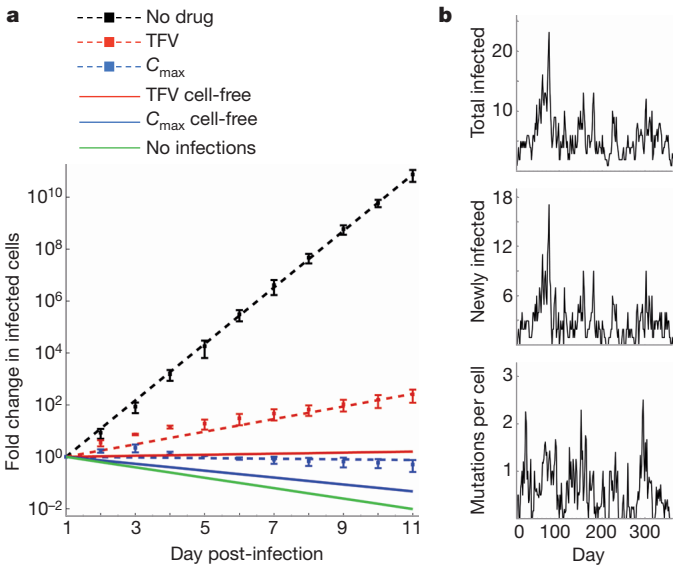



**Figure 3 | Co-culture infection dynamics.** **a**, Infection growth rate. Drug conditions were: no drug (black squares), 100 μM TFV (red squares) and $C_{max}$ (blue squares). Means ± s.d. of inter-day experiments ($n = 3$). Dashed lines represent fits of $I_k = I_0 R^k$ for each drug condition. Solid lines are predicted infection dynamics for infection occurring exclusively by cell-free virus in the presence of 100 μM TFV (red line), $C_{max}$ (blue line), or with no viral replication (green line). **b**, Simulation of the number of infected cells and mutations per cell with an input of one infected cell per virus cycle (Methods). $x$-axis is time, $y$-axis is number of total infected cells (top graph), newly infected cells (middle graph) or sum of mutations divided by the sum of total infected cells (bottom graph). Average number of mutations per cell over time is 0.6 ± 0.5 (mean ± s.d., $n = 215$).

## METHODS SUMMARY

**HIV infection at high and low *m*.** NL4-3YFP HIV stock at a 1:2,000 or 1:4 final dilution was added to MT-4 cells pre-incubated with TFV. Two days post-infection, the number of YFP-positive cells was determined by FACS. The multiplicity of infection was calculated as $m = -\ln p(0) = -\ln(1 - I_{1:2,000})$, where $I_{1:2,000}$ is the fraction of YFP-positive cells at the 1:2,000 dilution.

**Drug sensitivity of co-culture versus cell-free infections.** Donor cells were infected with NL4-3 strain HIV and incubated for two to three days. Infected donor cells or cell-free NL4-3 were then added to target cells. Two days after target cell infection, the number of infected cells was determined by FACS using intracellular p24 staining (PBMCs) or GFP expression (Rev-CEM cells). In all experiments, uninfected PBMC or MT-4mCherry cells were added to cell-free virus infections to keep total cell numbers equal on day 0.

**Infection growth rate.** Infection was initiated by adding infected Rev-CEM cells to uninfected Rev-CEM cells pre-incubated with drugs. Cells were passaged on each day following infected cell addition: infection with no drug was split 1:10 into fresh Rev-CEM cells. For 100 μM TFV or $C_{max}$, infected cells were split 0.6:1 with drug-containing medium. Cells remaining after split were used to quantify the fraction of infected cells by FACS. The fold change in infected cells on each day was calculated as $N_k D_k / N_0$, where $N_k$ is the fraction of infected cells on day $k$, $D_k$ is total dilution factor (split) up to day $k$ and $N_0$ is the fraction of infected cells on day 1.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 30 November 2010; accepted 1 July 2011.**
**Published online 17 August 2011.**

1. Pierson, T., McArthur, J. & Siliciano, R. F. Reservoirs for HIV-1: mechanisms for viral persistence in the presence of antiviral immune responses and antiretroviral therapy. *Annu. Rev. Immunol.* **18,** 665–708 (2000).
2. Jung, A. *et al.* Recombination: multiply infected spleen cells in HIV patients. *Nature* **418,** 144 (2002).
3. Dixit, N. M. & Perelson, A. S. HIV dynamics with multiple infections of target cells. *Proc. Natl Acad. Sci. USA* **102,** 8198–8203 (2005).
4. Dimitrov, D. S. *et al.* Quantitation of human immunodeficiency virus type 1 infection kinetics. *J. Virol.* **67,** 2182–2190 (1993).



©2011 Macmillan Publishers Limited. All rights reserved

5. Sattentau, Q. Avoiding the void: cell-to-cell spread of human viruses. *Nature Rev. Microbiol.* **6,** 815–826 (2008).
6. Martin, N. *et al.* Virological synapse-mediated spread of human immunodeficiency virus type 1 between T cells is sensitive to entry inhibition. *J. Virol.* **84,** 3516–3527 (2010).
7. Sourisseau, M., Sol-Foulon, N., Porrot, F., Blanchet, F. & Schwartz, O. Inefficient human immunodeficiency virus replication in mobile lymphocytes. *J. Virol.* **81,** 1000–1012 (2007).
8. Dang, Q. *et al.* Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. *Proc. Natl Acad. Sci. USA* **101,** 632–637 (2004).
9. Dixit, N. M. & Perelson, A. S. Multiplicity of human immunodeficiency virus infections in lymphoid tissue. *J. Virol.* **78,** 8942–8945 (2004).
10. Del Portillo, A. *et al.* Multiploid inheritance of HIV-1 during cell-to-cell infection. *J. Virol.* **85,** 7169–7176 (2011).
11. Buzón, M. J. *et al.* HIV-1 replication and immune dynamics are affected by raltegravir intensification of HAART-suppressed subjects. *Nature Med.* **16,** 460–465 (2010).
12. Chun, T. W. *et al.* Relationship between residual plasma viremia and the size of HIV proviral DNA reservoirs in infected individuals receiving effective antiretroviral therapy. *J. Infect. Dis.* **204,** 135–138 (2011).
13. Doitsh, G. *et al.* Abortive HIV infection mediates CD4 T cell depletion and inflammation in human lymphoid tissue. *Cell* **143,** 789–801 (2010).
14. Paredes, R. *et al.* Pre-existing minority drug-resistant HIV-1 variants, adherence, and risk of antiretroviral treatment failure. *J. Infect. Dis.* **201,** 662–671 (2010).
15. Levy, D. N., Aldrovandi, G. M., Kutsch, O. & Shaw, G. M. Dynamics of HIV-1 recombination in its natural target cells. *Proc. Natl Acad. Sci. USA* **101,** 4204–4209 (2004).
16. Gratton, S., Cheynier, R., Dumaurier, M. J., Oksenhendler, E. & Wain-Hobson, S. Highly restricted spread of HIV-1 and multiply infected cells within splenic germinal centers. *Proc. Natl Acad. Sci. USA* **97,** 14566–14571 (2000).
17. Collins, K. L., Chen, B. K., Kalams, S. A., Walker, B. D. & Baltimore, D. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. *Nature* **391,** 397–401 (1998).
18. Wu, Y., Beddall, M. H. & Marsh, J. W. Rev-dependent indicator T cell line. *Curr. HIV Res.* **5,** 394–402 (2007).
19. Chen, B. K., Gandhi, R. T. & Baltimore, D. CD4 down-modulation during infection of human T cells with human immunodeficiency virus type 1 involves independent activities of vpu, env, and nef. *J. Virol.* **70,** 6044–6053 (1996).
20. Ribeiro, R. M. *et al.* Estimation of the initial viral growth rate and basic reproductive number during acute HIV-1 infection. *J. Virol.* **84,** 6096–6102 (2010).
21. Nowak, M. A. & May, R. M. *Virus Dynamics: Mathematical Principles of Immunology and Virology* (Oxford Univ. Press, 2000).
22. Mathias, A. A. *et al.* Bioequivalence of efavirenz/emtricitabine/tenofovir disoproxil fumarate single-tablet regimen. *J. Acquir. Immune Defic. Syndr.* **46,** 167–173 (2007).
23. Bailey, J. R. *et al.* Residual human immunodeficiency virus type 1 viremia in some patients on antiretroviral therapy is dominated by a small number of invariant clones rarely found in circulating $CD4^+$ T cells. *J. Virol.* **80,** 6441–6457 (2006).
24. Kieffer, T. L. *et al.* Genotypic analysis of HIV-1 drug resistance at the limit of detection: virus production without evolution in treated adults with undetectable HIV loads. *J. Infect. Dis.* **189,** 1452–1465 (2004).
25. Cu-Uvin, S. *et al.* Genital tract HIV-1 RNA shedding among women with below detectable plasma viral load. *AIDS* **24,** 2489–2497 (2010).
26. North, T. W. *et al.* Viral sanctuaries during highly active antiretroviral therapy in a nonhuman primate model for AIDS. *J. Virol.* **84,** 2913–2922 (2010).
27. Rong, L. & Perelson, A. S. Modeling latently infected cell activation: viral and latent reservoir persistence, and viral blips in HIV-infected patients on potent therapy. *PLOS Comput. Biol.* **5,** e1000533 (2009).
28. Grossman, Z. *et al.* Ongoing HIV dissemination during HAART. *Nature Med.* **5,** 1099–1104 (1999).
29. Frenkel, L. M. *et al.* Multiple viral genetic analyses detect low-level human immunodeficiency virus type 1 replication during effective highly active antiretroviral therapy. *J. Virol.* **77,** 5721–5730 (2003).
30. Rhee, S. Y. *et al.* Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucleic Acids Res.* **31,** 298–303 (2003).


**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank B. K. Chen, A. Del Portillo, J. T. Schiffer, L. Corey, and G. Lustig for discussions. A.S. was supported by the Human Frontier Science Program Long Term Fellowship LT00946. J.T.K. was supported by the UCLA STAR fellowship and T32 AI089398. A.B.B. was supported by the amfAR Postdoctoral Research Fellowship 107756-47-RFVA. This work was supported by the Bill & Melinda Gates Foundation and by the National Institutes of Health (HHSN266200500035C) and a contract from the National Institute of Allergy and Infectious Diseases. We acknowledge the support of the UCLA CFAR Virology Core Lab (P01-AI-28697) and the UCSF-GIVI CFAR (P30-AI-27763).

**Author Contributions** A.S. and D.B. conceived the study. A.S. designed the research; A.S. and J.T.K. performed the experiments with support from A.B.B.; A.S. formulated the basic mathematical model and performed the numerical simulations; R.M., A.M. and E.D. added analytical insights and expanded the model to treat virus number as a random variable; A.S. and D.B. wrote the paper.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.B. (baltimo@caltech.edu).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Cells, viruses and drugs.** The following were obtained through the AIDS Research and Reference Reagent Program, National Institute of Allergy and Infectious Diseases, National Institutes of Health: Rev-CEM cells from Y. Wu and J. Marsh; MT-4 cells from D. Richman; HIV expression plasmid pNL4-3 from M. Martin; TFV; EFV. The NL4-3YFP molecular clone was a gift from D. Levy. Cell-free virus was produced by transfection of HEK293 cells with virus coding plasmid using Fugene 6 or Fugene HD (Roche). Supernatant containing shed virus was harvested after two days of incubation. Number of virus genomes of viral stock was determined using the RealTime HIV-1 Viral Load test (Abbott Diagnostics, Abbott Park Ill) and gag p24 content was determined by ELISA (Perkin-Elmer) at the the ARI-UCSF Laboratory of Clinical Virology. The MT-4mCherry cell line was created by infecting MT-4 cells with a pHAGE2 lentiviral vector expressing mCherry under the control of the EF1α promoter. To obtain a 0.99 fraction of mCherry-positive cells, MT-4mCherry cells were used fresh after lentiviral infection without a cycle of freezing and thawing, minimizing the number of population doublings and consequent decrease in the mCherry-positive fraction. Anonymous PBMCs or peripheral blood samples were provided by AllCells (PBMCs) or the UCLA Center for AIDS Research (CFAR) Virology Core Lab (peripheral blood). For whole blood, PBMCs were purified by Ficoll gradient using standard techniques. Purified PBMCs were activated with 5 μg ml$^{-1}$ PHA in the presence of 5 ng ml$^{-1}$ IL-2 for 1 (donors) or 3 (targets) days. All work was approved by the California Institute of Technology Institutional Biosafety Committee and Institutional Review Board exempt.

**NL4-3YFP infection at high and low *m*.** MT-4 cells were pre-incubated for 24 h with varying concentrations of TFV. NL4-3YFP stock was produced using transfection of HEK293 cells at 80% confluence with Fugene HD. Virus supernatant was collected 2 days post-transfection and added fresh to maximize the number of infectious units. Fresh virus stock was used at a 1:2,000 (low *m*) or 1:4 final dilution (high *m*). After 2 days incubation with virus, the number of YFP-positive MT-4 cells was quantified by flow cytometry by collecting $2 \times 10^5$ cells using a FACScaliber machine (Becton Dickenson). The multiplicity of infection was calculated using Poisson statistics: $m = -\ln p(0) = -\ln(1 - I_{1:2,000})$, where $p(0)$ is the fraction of YFP-negative cells, and $I_{1:2,000}$ is the fraction of YFP-positive cells at the 1:2,000 dilution.

**Comparison of co-culture and cell-free infections in PBMCs.** For PBMC infections, $1.5 \times 10^6$ PHA-activated HLA-A2-negative donor PBMCs at $10^6$ cells ml$^{-1}$ were either infected with 700 ng HIV (NL4-3 strain), or mock infected with the same volume of growth medium. Cells were then incubated for 2 days. Two days after donor-cell infection, PHA-activated HLA-A2-positive PBMC target cells at $10^6$ cells ml$^{-1}$ were either treated with no drug or 10 μM TFV. The stock of target cells with or without drug was then split into wells at $10^6$ cells well$^{-1}$ and incubated for 4 h. After target cell incubation, HLA-A2-negative donor PBMCs were washed, counted, diluted to $10^6$ cells ml$^{-1}$ and added to target cells at an approximately 1:10 donor:target ratio as follows. For cell-free infection, each well received 100 μl mock-infected HLA-A2-negative donor PBMCs and 150 μl (250 ng) cell-free NL4-3. For co-culture infection, each well received 100 μl infected HLA-A2-negative donor PBMCs and 150 μl growth medium. One day after target-cell infection, cell aggregates were broken up by repeated pipetting, and cells split 1:2 with fresh growth medium containing the corresponding drug concentration. Two days after target-cell infection, the number of infected target cells was determined: cells were stained with PE-conjugated anti-HLA-A2 antibody (BD Biosciences or Biolegend), fixed and permeabilized (Cytofix/cytoperm kit, BD Biosciences), then stained with intracellular FITC-conjugated anti-HIV p24 antibody (clone KC57, Coulter Corporation) according to the Cytofix/cytoperm kit protocol. The fraction of infected target cells was quantified by FACS as HLA-A2, p24 double-positive cells. We observed that PBMCs were infected best when fresh, and use of previously frozen material or cells whose processing was delayed substantially reduced both cell-free and co-culture infections.

**Comparison of co-culture and cell-free infections using Rev-CEM cells.** We infected Rev-CEM target cells either by co-culture with MT-4mCherry donor cells or cell-free virus. MT-4mCherry donor cells at $4 \times 10^5$ cells ml$^{-1}$ were infected with 300 ng ml$^{-1}$ p24 NL4-3, or mock infected with the same volume of growth medium. Donor cells were then incubated for 3 days. Two days after donor-cell infection and one day before target-cell infection, Rev-CEM target cells at $8 \times 10^5$ cells ml$^{-1}$ were treated with no drug, TFV, or EFV. The stock of target cells with or without drug was then split into wells at $1.6 \times 10^6$ cells well$^{-1}$ and incubated for 24 h. Three days after donor-cell infection, MT-4mCherry donor cells were washed, counted, diluted to $3 \times 10^5$ cells ml$^{-1}$ and added at an approximately 1:100 donor:target ratio as follows. For cell-free infection, each well received 100 μl mock-infected MT-4mCherry donor cells and 600 μl (1 μg) cell-free NL4-3. For co-culture infection, each well received 100 μl infected MT-4mCherry donor cells and 600 μl growth medium. One day after target-cell infection, cell aggregates were broken up by repeated pipetting, and cells split 1:2 with fresh growth medium containing the corresponding drug concentration. Two days after target-cell infection, the number of infected target cells were quantified by FACS by mCherry and GFP fluorescence. Infected target cells were gated as positive for GFP, and negative for mCherry, thereby excluding uninfected Rev-CEM cells (GFP negative), MT-4mCherry cells (GFP negative, mCherry positive), and fusions between MT-4mCherry and Rev-CEM cells (GFP positive, mCherry positive). The fraction of MT-4mCherry donors was 1% on day 0 for both mock-infected and infected cells, but decreased for infected MT-4mCherry cells by the end of the target-cell infection, probably owing to the cytotoxicity of infection. To ensure that the low numbers of infected target cells gave repeatable results, we averaged consecutive independent inter-day experiments.

**Infection growth rate.** To initiate infection, Rev-CEM cells at $4 \times 10^5$ cells ml$^{-1}$ were infected with 300 ng ml$^{-1}$ NL4-3 in the absence of drugs and incubated for three days. Two days post-infection, uninfected Rev-CEM cells at $8 \times 10^5$ cells ml$^{-1}$ were pre-treated with no drug, 100 μM TFV, or a combination of EFV, TFV and FTC at their clinical maximum plasma concentrations ($C_{max}$: 10 μM EFV, 2 μM TFV and 10 μM FTC). Three days after the initial infection with cell-free virus, the infected Rev-CEM cells were washed and added to a final fraction of 0.2% GFP-expressing donor cells to the uninfected Rev-CEM cells incubated with no drug, TFV or $C_{max}$. On each day after infected donor addition, cell aggregates were broken up by gentle repeated pipetting and cells split. Infection conditions were calibrated so that the number of uninfected target cells would not be limiting and infection would not interfere with proliferation of uninfected cells. Infection was therefore kept below ~0.5% GFP-expressing infected Rev-CEM cells. The daily cell dilution was calibrated to keep this steady state of infected cells: the sample with no drug was split 1:10 or 1:20 into fresh Rev-CEM cells in a new well. For 100 μM TFV or $C_{max}$ drug concentrations, infected cells were split 0.6:1 with drug-containing medium into a new well. Cells remaining after cell split were used to quantify the fraction of infected cells by FACS ($5 \times 10^5$ collected per sample). The fold change in infected cells on each day was calculated as $N_k D_k / N_0$, where $N_k$ is the fraction of infected cells on day $k$, $D_k$ is the total dilution factor up to day $k$ and $N_0$ is the fraction of infected cells on day 1. The drug effect on a single round of cell-free infection for 100 μM TFV or $C_{max}$ was measured at the same time as the infection growth rate to prevent differences in drug stock batch or cells.

**Stochastic simulation of the number of infected cells and mutations.** The purpose of the simulation was to determine the sum of total infected cells, newly infected cells, and mutations at each virus cycle (measured as 1.7 days (Supplementary Fig. 2)) from overlapping infection chains. A new infection chain was initiated each virus cycle with an input of one infected cell. The number of infected cells in cycle $k+1$ generated by infected cell $j$ in cycle $k$ was an integer ${I_{k+1}}^j = x_1 + x_2$, where $x_1$ was a random number from a Poisson distribution with an average $\mu_1$ defined by the measured infected-cell half-life $\mu_1 = 2^{-\tau/t1/2} = 0.46$ (Supplementary Fig. 13 and Supplementary Theory, section 3), and $x_2$ was a random number from a Poisson distribution with an average $\mu_2$ defined by $\mu_2 = R_{C_{max}} - \mu_1 = 0.49$ (Supplementary Theory, section 3). Given an outcome of $N$ infected cells in cycle $k$, the number of total infected cells in virus cycle $k+1$ in the infection chain was $\sum_1^N (x_1^j + x_2^j)$, of which the number of newly infected cells was $\sum_1^N x_2^j$. A new infection chain from an input of one infected cell was generated every virus cycle. Therefore, infection chains overlapped, and the total output number of infected cells in virus cycle $k+1$ was a sum of infected cells at that virus cycle from all $M$ infection chains: $\sum_1^M \sum_1^N (x_1^j + x_2^j)$. The number of newly infected cells was $\sum_1^M \sum_1^N x_2^j$. If a new infection occurred, the probability of mutation occurring at any one of the $10^4$ nucleotides of the HIV genome was $1 - (1 - 3.4 \times 10^{-5})^{10^4} = 0.29$, where $3.4 \times 10^{-5}$ is the per-base probability of mutation for the HIV reverse transcriptase, and $(1 - 3.4 \times 10^{-5})^{10^4}$ is the probability that no mutations occur during a single reverse transcription event. As a simplifying assumption, no fitness benefit or cost was assigned to individual mutations. Therefore, $R_{C_{max}}$ did not change during the course of the simulation. Surviving cells carried over their mutations to the next generation, and newly infected cells carried over mutations from the infected donor cells, in addition to any mutations generated during the infection process.

©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10363

# Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase

Christian Frezza[1], Liang Zheng[1], Ori Folger[2], Kartik N. Rajagopalan[3], Elaine D. MacKenzie[1], Livnat Jerby[2], Massimo Micaroni[4], Barbara Chaneton[1], Julie Adam[5], Ann Hedley[1], Gabriela Kalna[1], Ian P. M. Tomlinson[6], Patrick J. Pollard[5], Dave G. Watson[7], Ralph J. Deberardinis[3], Tomer Shlomi[8]*, Eytan Ruppin[2,9]* & Eyal Gottlieb[1]

**Fumarate hydratase (FH) is an enzyme of the tricarboxylic acid cycle (TCA cycle) that catalyses the hydration of fumarate into malate. Germline mutations of *FH* are responsible for hereditary leiomyomatosis and renal-cell cancer (HLRCC)[1]. It has previously been demonstrated that the absence of *FH* leads to the accumulation of fumarate, which activates hypoxia-inducible factors (HIFs) at normal oxygen tensions[2–4]. However, so far no mechanism that explains the ability of cells to survive without a functional TCA cycle has been provided. Here we use newly characterized genetically modified kidney mouse cells in which *Fh1* has been deleted, and apply a newly developed computer model of the metabolism of these cells to predict and experimentally validate a linear metabolic pathway beginning with glutamine uptake and ending with bilirubin excretion from *Fh1*-deficient cells. This pathway, which involves the biosynthesis and degradation of haem, enables *Fh1*-deficient cells to use the accumulated TCA cycle metabolites and permits partial mitochondrial NADH production. We predicted and confirmed that targeting this pathway would render *Fh1*-deficient cells non-viable, while sparing wild-type *Fh1*-containing cells. This work goes beyond identifying a metabolic pathway that is induced in *Fh1*-deficient cells to demonstrate that inhibition of haem oxygenation is synthetically lethal when combined with *Fh1* deficiency, providing a new potential target for treating HLRCC patients.**

To study metabolic adaptation of *FH*-deficient cells we first generated immortalized kidney cells from mice homozygous for a conditionally targeted *Fh1* allele[3]. These cells contain LoxP sites flanking exons 3 and 4 ($Fh1^{fl/fl}$) (Fig. 1a and Supplementary Fig. 2a) and express normal levels of Fh1 protein (Fig. 1b). To generate $Fh1^{-/-}$ cells the $Fh1^{fl/fl}$ cells were infected with recombinant adenovirus expressing Cre recombinase and two clones (CL1 and CL19) were selected from the infected pools. Both clones were genetically confirmed to contain homozygous *Fh1*-deleted (knockout) alleles (Fig. 1a) and did not express Fh1 protein (Fig. 1b).

To assess the biochemical consequences of *Fh1* deletion on TCA cycle function, the intracellular levels of several TCA cycle metabolites were measured by gas chromatography–mass spectrometry (GC–MS) (Fig. 1c). $Fh1^{-/-}$ cells accumulate substantial amounts of fumarate, which reached levels more than a 100-fold that of $Fh1^{fl/fl}$ cells. In addition, a sevenfold increase in succinate levels and a marked decrease of malate and citrate were detected (Fig. 1d). In cancer cells, the TCA cycle is largely supplied through the metabolism of glucose and glutamine[5]. To investigate the effects of *Fh1* deletion on carbon supply to the TCA cycle, cells were cultured in medium containing $^{13}$C-labelled glucose and unlabelled glutamine, or vice versa, and $^{13}$C-enrichment of fumarate and succinate was analysed. When $^{13}$C-glucose was used, the majority of fumarate was unlabelled (m+0), indicating that glucose is a minor source of carbon for the TCA cycle in both $Fh1^{fl/fl}$ and $Fh1^{-/-}$ cells (Fig. 1e). On the other hand, when cells were incubated with $^{13}$C-glutamine most of the fumarate was labelled (Fig. 1f). In $Fh1^{-/-}$ cells cultured with uniformly labelled glutamine (Fig. 1f), the vast majority of the labelled fumarate contained all four carbon atoms derived from glutamine (m+4). By contrast, in $Fh1^{fl/fl}$ cells, the fumarate pool contained substantial fractions of molecules with fewer than four $^{13}$C atoms due to processing of fumarate beyond the Fh1 step. The lack of these products in $Fh1^{-/-}$ cells confirms that no accessory pathway exists in these cells to circumvent the loss of Fh1 enzymatic activity, indicating a true blockade of the cycle. Notably, the $^{13}$C-enrichment profiles of

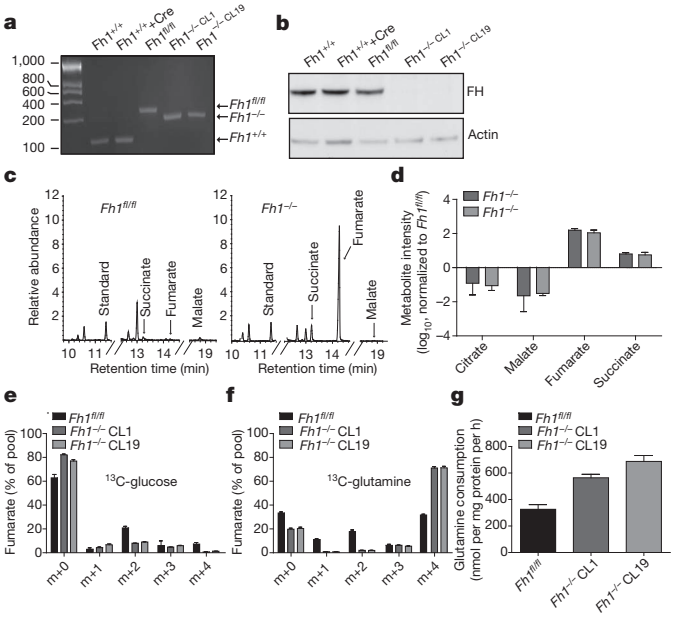



**Figure 1 | Characterization of *Fh1*-deficient cells. a**, PCR analysis of genomic DNA of epithelial kidney cells of the indicated genotype. Numbers indicate base pairs. **b**, Fh1 protein levels of the indicated cells were detected by western blot analysis. **c**, Representative chromatogram of a GC–MS analysis performed from the indicated cell lines. Annotated metabolites are indicated by the arrows. **d**, Fold change of the indicated TCA cycle metabolites in $Fh1^{-/-}$ compared to control $Fh1^{fl/fl}$ cells ($n = 3$). **e**, **f**, Isotopomer distribution of intracellular fumarate. Cells were incubated with either $^{13}$C-glucose (**e**) or $^{13}$C-glutamine (**f**) and fumarate isotopomers (with the indicated mass shift) were analysed ($n = 3$). **g**, Glutamine consumption rate of the indicated cell lines ($n = 3$). All data are presented as mean ± s.e.m.

[1]Cancer Research UK, Beatson Institute for Cancer Research, Switchback Road, Glasgow G61 1BD, UK. [2]The Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel. [3]Department of Pediatrics and McDermott Center for Human Growth and Development, University of Texas–Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, Texas 75390-9063, USA. [4]The University of Queensland, Institute for Molecular Bioscience, St Lucia, Brisbane, Queensland 4072, Australia. [5]Henry Wellcome Building for Molecular Physiology, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. [6]Molecular and Population Genetics Laboratory, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK. [7]Strathclyde Institute of Pharmaceutical and Biomedical Sciences, University of Strathclyde, 27 Taylor Street, Glasgow G4 0NR, UK. [8]Computer Science Department, Technion, Israel Institute of Technology, Haifa, 32000, Israel. [9]The Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel.

*These authors contributed equally to this work.



©2011 Macmillan Publishers Limited. All rights reserved

succinate perfectly matched fumarate results, underlining the presence of a flux of carbons from glutamine to fumarate via succinate. In addition, no reverse flux of TCA cycle metabolites was detected (Supplementary Fig. 2b, c). (See Supplementary Fig. 1 for a diagram summarizing these results.) Consistent with the increased contribution of glutamine to the pool of TCA cycle intermediates, $Fh1^{-/-}$ cells used significantly more glutamine than $Fh1^{fl/fl}$ cells (Fig. 1g).

Because the TCA cycle is a major source for mitochondrial NADH, the truncation of the TCA cycle may have severe bioenergetic outcomes. Indeed, steady-state levels of mitochondrial NADH in $Fh1^{-/-}$ cells were lower than those of control cells (Fig. 2a; untreated and Supplementary Fig. 2d for quantification). However, $Fh1^{-/-}$ cells still generate significant amounts of NADH and retained normal mitochondrial membrane potential (Fig. 2a). In addition, depolarizing the mitochondria by using the protonophore carbonyl cyanide *m*-chlorophenyl hydrazone (CCCP) decreased mitochondrial NADH, whereas blocking NADH oxidation with rotenone, an inhibitor of NADH:ubiquinone oxidoreductase, increased mitochondrial NADH in both control and *Fh1*-deficient cells. These results indicate that despite having a truncated TCA cycle, *Fh1*-deficient cells retain significant mitochondrial bioenergetic activity and are capable of generating and oxidizing NADH. Because NADH oxidation in the mitochondria fuels the respiratory chain, the oxygen consumption rate was analysed. $Fh1^{-/-}$ cells showed a significant reduction of their cyanide-dependent respiratory rate (Fig. 2b) when compared to wild-type cells. The decrease in respiration was not due to a general reduction in mitochondrial number as was assessed both microscopically (Fig. 2a) and by mitochondrial DNA quantification (Supplementary Fig. 2e). Furthermore, and as was recently reported for HLRCC (Online Mendelian Inheritance in Man accession number 605839) patient-derived cells[6], although mitochondria of $Fh1^{-/-}$ cells appear larger under electron microscopy they do retain some organized cristae structures (Supplementary Fig. 2f). The

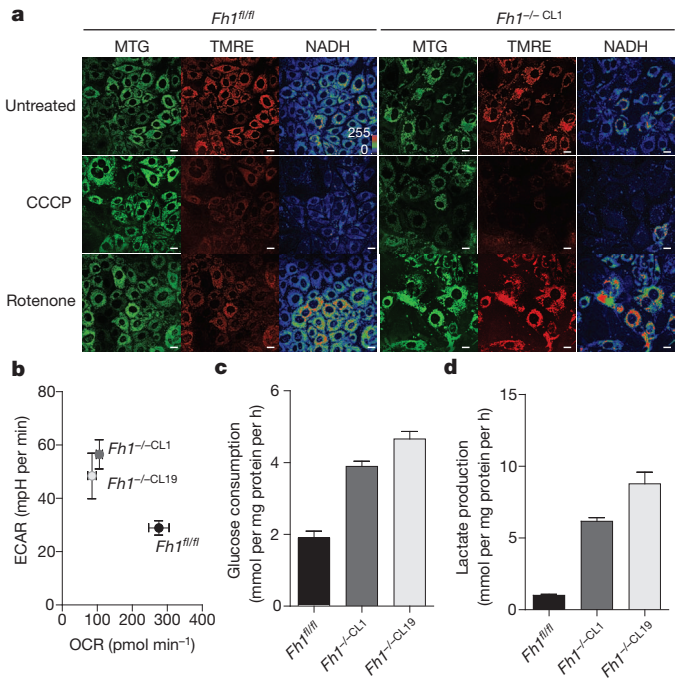



**Figure 2 | Metabolic analyses of $Fh1^{-/-}$ cells.** **a**, Confocal microscopy was used to assess mitochondrial membrane potential and mitochondrial NADH levels in the indicated cell lines. Where indicated, cells were incubated with rotenone or CCCP before visualization. Scale bar: 50 μm. MTG, MitoTracker green; TMRE, tetramethyl rhodamine ethyl ester. **b**, Representative graph of oxygen consumption rate (OCR) versus extracellular acidification rate (ECAR). Data are presented as mean ± s.d. ($n = 6$). **c**, **d**, Glucose consumption (**c**) and lactate production (**d**) of the indicated cell lines. Data are presented as mean ± s.e.m. ($n = 3$).

decrease in respiration of $Fh1^{-/-}$ cells was associated with an increase in glycolysis as indicated initially by the increased extracellular acidification rate (Fig. 2b) and further confirmed by the increase in glucose consumption and lactate production (Fig. 2c, d).

Metabolic networks are highly interconnected, making them resistant to most single gene perturbations[7]. To elucidate the metabolic changes that enable the survival of $Fh1^{-/-}$ cells despite having a truncated TCA cycle, we took a systemic approach based on flux balance analysis (FBA) to obtain a genome-scale model of cellular metabolism, as very commonly used in modelling of microbial metabolism[8]. We adopted an *in silico* modelling approach for cancer metabolism[9] (see Supplementary Information for details) to predict genes for which elimination together with *Fh1* would selectively affect the growth ability of $Fh1^{-/-}$ cells without affecting the wild-type $Fh1^{fl/fl}$ cells (predicted to be synthetic lethal with *Fh1*). This modelling involved the reconstruction of genome-scale metabolic network models of the metabolism of $Fh1^{fl/fl}$ and $Fh1^{-/-}$ cells, using a generic computational method for building metabolic models of specific human tissues[10]. The reconstruction of the models began with the identification of core sets of metabolic-enzyme-coding genes that are expressed in $Fh1^{fl/fl}$ and $Fh1^{-/-}$ cells, based on accompanying microarray gene expression analysis (Supplementary Fig. 3), added to a set of metabolic genes that are highly expressed across many cancer cell lines[11]. Then, for each cell line, a minimal, consistent metabolic network that includes the corresponding core reactions set was extracted from an existing generic human model[12], such that each of the core reactions can potentially carry a stoichiometrically balanced metabolic flux (Supplementary Table 1). As conventionally done in FBA modelling[13], cellular growth requirements are modelled via the addition of a growth reaction, which represents the steady-state consumption of essential biomass constituents required for cellular proliferation, based on previous knowledge of their relative concentrations. The resulting metabolic models were used to predict the consequences of gene knockouts based on their effect on reducing biomass production rate, considering the rate of glucose, glutamine and oxygen uptake as well as lactate secretion measured in each of the cell lines (Supplementary Table 2).

Overall, 24 reactions were predicted by this analysis to be synthetic lethal with *Fh1*, as their deletion is predicted to affect growth only in the $Fh1^{-/-}$ metabolic network model (Fig. 3a and Supplementary Table 3). Of these, 18 belong to a linear pathway of haem metabolism, which was also predicted by the model to have increased flux in $Fh1^{-/-}$ cells. Haem biosynthesis is a cataplerotic pathway that uses TCA-cycle-derived carbon to generate haem, whereas the haem degradation pathway generates bilirubin that can be removed from the cells. Thus, haem biosynthesis and degradation would generate a linear pathway starting with glutamine uptake and ending with bilirubin excretion. This biochemical escape valve would allow some generation of mitochondrial NADH (Supplementary Fig. 1). Interestingly, the transcriptomic analysis revealed that three of the predicted synthetic lethal genes were also upregulated in $Fh1^{-/-}$ cells (Fig. 3b): haem oxygenase 1 (*Hmox1*; but not *Hmox2*), which catabolizes haem to biliverdin; biliverdin reductase B (*Blvrb*), which catabolises biliverdin to bilirubin; and UDP-glucoronyl transferase 2b34 (*Ugt2b34*), which facilitates bilirubin excretion. The expression levels of these genes were validated (Fig. 3c, d) and the increase in Hmox1 protein levels was also confirmed (Fig. 3e). None of these transcriptional changes could be simply attributed to the expression of Cre recombinase, because they were not observed in the Cre-infected $Fh1^{+/+}$ cells, which do not express the conditional (LoxP) *Fh1* alleles (Fig. 3c, d). Notably, although these three genes exhibit significantly increased expression in $Fh1^{-/-}$ cells, the potential role of this metabolic pathway in the survival of the *Fh1*-deficient cells could not have been predicted solely by its gene expression response to *Fh1* knockout, as many additional pathways show a similar increase in expression (Supplementary Fig. 3b).

To validate the activation of the haem pathway in $Fh1^{-/-}$ cells, as predicted by the FBA model, we measured the levels of excreted bilirubin in the media using liquid chromatography–mass spectrometry



©2011 Macmillan Publishers Limited. All rights reserved

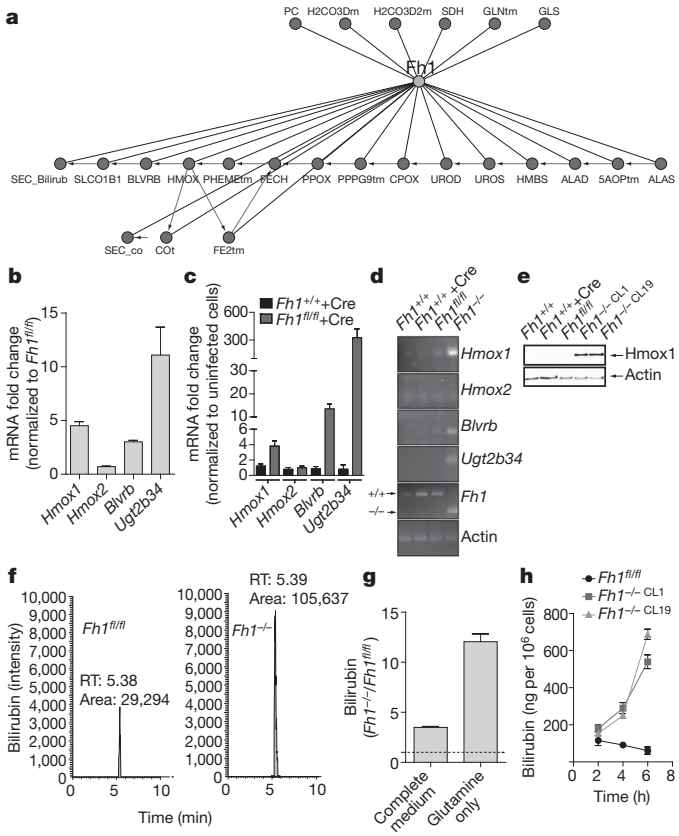



**Figure 3 | The haem pathway is predicted to be synthetically lethal with *Fh1*.** **a**, Synthetic lethal pairs with *FH* were predicted using the FBA metabolic modelling. The haem synthesis and degradation pathway (bottom) was highly represented. 5AOPtm, aminolevulinic acid (ALA) transport; ALAD, ALA dehydratase; ALAS, ALA synthase; BLVRB, biliverdin reductase; COt, carbon monoxide (CO) transport; CPOX, coproporphyrinogen oxidase; FE2tm, iron transport; FECH, ferrochelatase; GLNtm, glutamine transport; GLS, glutaminase; H2CO3Dm/D2m, carbonate transport; HMBS, hydroxymethylbilane synthase; HMOX, haem oxygenase; PC, pyruvate carboylase; PHEMEtm, haem transporter; PPOX, protoporphyrinogen oxidase; PPPG9tm, porphyrin transporter; SDH, succinate dehydrogenase; SEC_Bilirub, bilirubin excretion; SEC_co, CO excretion; SLCO1B1, bilirubin transporter; UROD, uroporphyrinogen (URO) decarbolxylase; UROS, URO synthase. **b**, Transcriptomic analysis of $Fh1^{-/-}$ cells revealed an increased expression of three genes involved with the haem degradation pathway: *Hmox1*, *Blvrb* and UDP-glucoronyl transferase 2b34 (*Ugt2b34*). Data are presented as mean ± s.d. ($n = 2$). **c**, **d**, The expression levels of the genes indicated in **b** were verified using qPCR (**c**) and end-point RT–PCR (**d**). Data are presented as mean ± s.e.m. ($n = 3$). **e**, Hmox1 protein levels in the indicated cell lines were detected by western blot analysis. **f**, Representative LC–MS chromatogram of bilirubin excreted by the indicated cell lines after 24 h in complete medium. RT, retention time (min); area is expressed as arbitrary units and was calculated by the integration of the bilirubin peak. **g**, Analysis of bilirubin excretion from the indicated cell lines grown either in complete medium or in the absence of glucose for 6 h. Data are presented as mean ± s.e.m. ($n = 3$). **h**, Kinetics of bilirubin excretion from the indicated cell lines incubated in glucose-free medium. Data are presented as mean ± s.d. ($n = 2$).

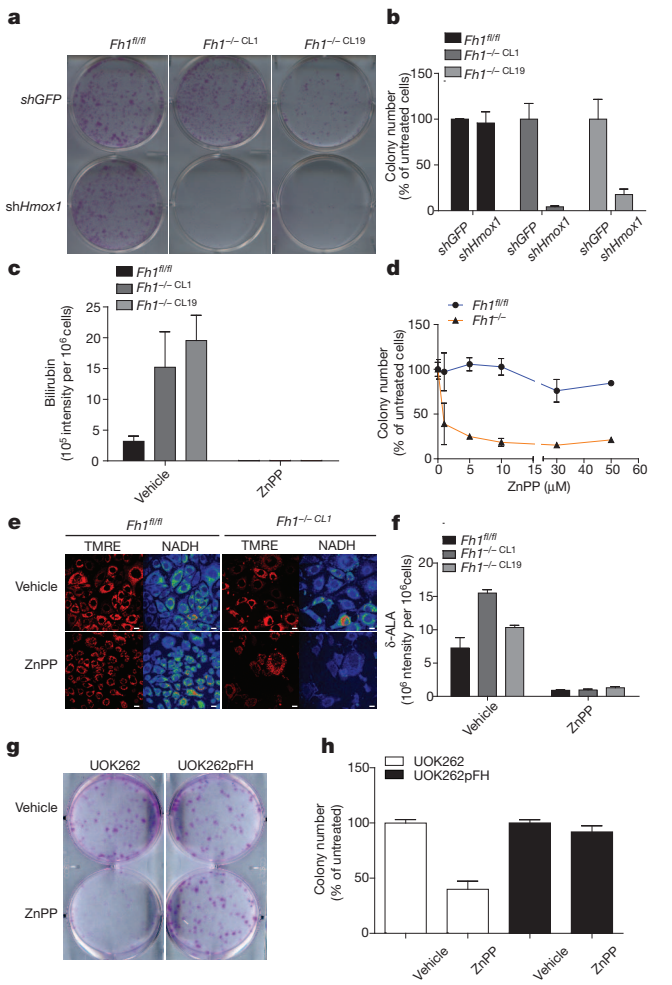



**Figure 4 | Targeting the haem degradation pathway is lethal for *Fh1*-deficient cells.** **a**, Representative images of a colony assay of cells infected with lentiviruses expressing either a non-targeting shRNA sequence (*shGFP*) or with an shRNA targeting *Hmox1*. **b**, Quantification of the clonogenic assay described in **a**. **c**, Quantification of bilirubin excretion before and after treatment with 10 μmol $l^{-1}$ ZnPP for 24 h. **d**, Clonogenic survival assay of wild-type or *Fh1*-deficient cells treated with the indicated concentration of ZnPP. **e**, The effect of ZnPP (treated as in **c**) on mitochondrial NADH was analysed microscopically. **f**, The effect of ZnPP (treated as in **c**) on intracellular ALA levels was measured by LC–MS. **g**, **h**, Representative images (**g**) and colony quantification (**h**) of a clonogenic survival assay of UOK262 and UOK262pFH cells treated with 1 μmol $l^{-1}$ ZnPP for 4 days. All the data are presented as mean ± s.e.m., ($n = 3$).

(LC–MS) (Fig. 3f). In agreement with our model, the levels of excreted bilirubin were higher in *Fh1*-deficient cells (Fig. 3f–h). Importantly, excretion of bilirubin in $Fh1^{-/-}$ cells was further increased when cells were transiently deprived of glucose (Fig. 3g), underpinning their increased dependency on glutamine as a mitochondrial carbon source during glucose deprivation.

In support of these findings, analysis of a published transcriptomic data set of human uterine fibroids[14] revealed significant upregulation of *HMOX1* (but not *HMOX2*) in *FH*-mutant fibroids (Supplementary Fig. 4). These results prompted us to validate the predicted synthetic lethality between *Fh1* and *Hmox1* in our mouse kidney cell model. First, *Hmox1* was efficiently silenced by lentiviral infection of short hairpin (sh)RNAs (*shHmox1* 1–3) (Supplementary Fig. 5a). All three shRNAs demonstrated significant reduction in growth of *Fh1*-deficient cells, whereas there was minimal to no effect on the control cells (Supplementary Fig. 5b). The effect of the most efficient shRNA (*shHmox1* 1) on long-term colony survival was further studied and it resulted in a marked decrease in colony number in $Fh1^{-/-}$ cells but not in wild-type cells (Fig. 4a, b). Importantly, the surviving *shHmox1*-infected $Fh1^{-/-}$ cells escaped silencing and their *Hmox1* expression levels reverted to that of mock-infected cells (Supplementary Fig. 5c).

For further confirmation of the above findings, we used zinc protoporphyrin (ZnPP), a Hmox inhibitor[15]. First, the efficacy of Hmox inhibition by ZnPP was tested by measuring bilirubin excretion. In both wild-type and *Fh1*-deficient cells bilirubin excretion was completely blocked by ZnPP (Fig. 4c). Whereas acute treatment with ZnPP had no profound effect on wild-type cells, it decreased the growth of $Fh1^{-/-}$ cells (Supplementary Fig. 5d). More importantly, similarly to



©2011 Macmillan Publishers Limited. All rights reserved

*Hmox1* silencing, long-term treatment with ZnPP abolished colony formation of $Fh1^{-/-}$ cells but had very little effect on wild-type cells (Fig. 4d). Furthermore, treatment of *Fh1*-deficient cells with ZnPP resulted in a decrease in mitochondrial NADH, whereas only slight changes were observed in control cells (Fig. 4e). These observations further elucidate the bioenergetic support that the glutamine–bilirubin linear pathway provides to *Fh1*-deficient cells.

δ-Aminolevulinic acid (ALA) synthase (ALAS), the entry point and rate-limiting step of haem biosynthesis (Supplementary Fig. 1), is tightly (negatively) controlled by haem and protoporphyrin to prevent the accumulation of porphyrins and its toxic consequences (porphyria)[16]. Therefore, it is likely that by blocking haem degradation, ZnPP can also inhibit the activity of ALAS. Indeed, the intracellular levels of ALA were decreased in ZnPP-treated cells (Fig. 4f), indicating that ZnPP has profound effects on the entire haem pathway beyond the inhibition of Hmox. To confirm further the importance of the cataplerotic role of heam biosynthesis in supporting *Fh1*-deficient cells, we also inhibited ALAS using hemin, an approved drug used for acute porphyria[17]. Similarly to Hmox inhibition, the inhibition of ALAS had no effect on control cells, whereas *Fh1*-deficient cells were sensitive to hemin treatment (Supplementary Fig. 5e).

Finally, to extend our findings to a human model of *FH* deficiency, we studied the recently described HLRCC-derived renal cancer cell line UOK262 (ref. 6). These cells are pseudo-hypoxic due to the accumulation of fumarate[18]. UOK262 cells were reintroduced with V5-tagged FH protein which was correctly localized to the mitochondria, improved oxygen consumption and restored low levels of fumarate (Supplementary Fig. 6a–c). Importantly, restoring TCA function to the *FH*-deficient UOK262 cells also markedly reduced bilirubin secretion and rendered the cells less sensitive to the HMOX and ALAS inhibitors ZnPP and hemin, respectively (Fig. 4g, h and Supplementary Fig. 6d, e). Importantly, the sensitivity of the FH-deficient UOK262 cells to ZnPP is independent on HIF activity (Supplementary Fig. 6f–h).

Taken together, the results of this work demonstrate that the inhibition of the haem biosynthesis/degradation pathway in general, and Hmox in particular, is synthetically lethal with *Fh1* both in mouse and in human cell lines. Hence, its inhibition would provide a valid therapeutic window for the treatment of HLRCC patients as it would specifically target *FH*-deficient cells while sparing normal tissues, thus reducing nonspecific toxic effects.

## METHODS SUMMARY

Epithelial kidney cells were obtained from $Fh1^{fl/fl}$ mice following the method described previously[19]. The $Fh1^{-/-}$ cell line was generated by infection with AhCRE adenoviral particles. The genotype was assessed by PCR on genomic DNA using the primers described in the Methods. TCA metabolites and their isotopomer enrichment analyses were performed by GC–MS after incubating the cells with either U-$^{13}$C-glucose or U-$^{13}$C-glutamine. Mitochondrial potential and NADH were analysed by confocal microscopy using the potentiometric probe TMRE and NADH autofluorescence, respectively. Glucose and glutamine consumption and lactate secretion were measured using the BioProfile Basic4 Analyser. Oxygen consumption rate and extracellular acidification rate were measured using the Seahorse XF24. Expression levels of the genes in the haem pathway were determined by qPCR and validated using semi-quantitative RT–PCR as detailed in the Methods. Bilirubin, ALA, citrate and aconitate levels were measured by LC–MS. Sulphorhodamine B was used to stain colonies. Colonies were counted using ImageJ software.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 13 July 2010; accepted 11 July 2011.**
**Published online 17 August 2011.**

1. Tomlinson, I. P. *et al.* Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nature Genet.* **30,** 406–410 (2002).
2. Isaacs, J. S. *et al.* HIF overexpression correlates with biallelic loss of fumarate hydratase in renal cancer: novel role of fumarate in regulation of HIF stability. *Cancer Cell* **8,** 143–153 (2005).
3. Pollard, P. J. *et al.* Targeted inactivation of fh1 causes proliferative renal cyst development and activation of the hypoxia pathway. *Cancer Cell* **11,** 311–319 (2007).
4. Frezza, C., Pollard, P. J. & Gottlieb, E. Inborn and acquired metabolic defects in cancer. *J. Mol. Med.* **89,** 213–220 (2011).
5. DeBerardinis, R. J. *et al.* Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proc. Natl Acad. Sci. USA* **104,** 19345–19350 (2007).
6. Yang, Y. *et al.* UOK 262 cell line, fumarate hydratase deficient *(FH$^{-}$/FH$^{-}$)* hereditary leiomyomatosis renal cell carcinoma: *in vitro* and *in vivo* model of an aberrant energy metabolic pathway in human cancer. *Cancer Genet. Cytogenet.* **196,** 45–55 (2010).
7. Deutscher, D., Meilijson, I., Kupiec, M. & Ruppin, E. Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nature Genet.* **38,** 993–998 (2006).
8. Price, N. D., Papin, J. A., Schilling, C. H. & Palsson, B. O. Genome-scale microbial *in silico* models: the constraints-based approach. *Trends Biotechnol.* **21,** 162–169 (2003).
9. Folger, O. *et al.* Predicting selective drug targets in cancer through metabolic networks. *Mol. Syst. Biol.* **7,** 501 (2011).
10. Jerby, L., Shlomi, T. & Ruppin, E. Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. *Mol. Syst. Biol.* **6,** 401 (2010).
11. Lee, J. K. *et al.* A strategy for predicting the chemosensitivity of human cancers and its application to drug discovery. *Proc. Natl Acad. Sci. USA* **104,** 13086–13091 (2007).
12. Duarte, N. C. *et al.* Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl Acad. Sci. USA* **104,** 1777–1782 (2007).
13. Orth, J. D., Thiele, I. & Palsson, B. O. What is flux balance analysis? *Nature Biotechnol.* **28,** 245–248 (2010).
14. Vanharanta, S. *et al.* Distinct expression profile in fumarate-hydratase-deficient uterine fibroids. *Hum. Mol. Genet.* **15,** 97–103 (2006).
15. Labbe, R. F., Vreman, H. J. & Stevenson, D. K. Zinc protoporphyrin: A metabolite with a mission. *Clin. Chem.* **45,** 2060–2072 (1999).
16. Schoenfeld, N. *et al.* The effects of metalloporphyrins, porphyrins and metals on the activity of delta-aminolevulinic acid synthase in monolayers of chick embryo liver cells. *Biochem. Pharmacol.* **33,** 2783–2788 (1984).
17. Anderson, K. E. & Collins, S. Open-label study of hemin for acute porphyria: clinical practice implications. *Am. J. Med.* **119,** e19–24 (2006).
18. Sudarshan, S. *et al.* Fumarate hydratase deficiency in renal cancer induces glycolytic addiction and hypoxia-inducible transcription factor 1α stabilization by glucose-dependent generation of reactive oxygen species. *Mol. Cell. Biol.* **29,** 4080–4090 (2009).
19. Mathew, R., Degenhardt, K., Haramaty, L., Karp, C. M. & White, E. Immortalized mouse epithelial cell models to study the role of apoptosis in cancer. *Methods Enzymol.* **446,** 77–106 (2008).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** This work was supported by Cancer Research UK. We would like to acknowledge the Patterson Institute for Cancer Research for the Exon Array analysis. C.F. is supported by an EMBO long term fellowship (ALTF330). P.J.P. is in receipt of a Beit Memorial Fellowship funded by the Wellcome Trust (WT091112MA). D.G.W. and L.Z. would like to acknowledge SULSA (Scottish Universities Life Science Alliance). R.J.D. and K.N.R. are supported by the NIH (DK072565-05) and the Cancer Prevention and Research Institute of Texas (HIRP100437-01). O.F., T.S. and E.R. are supported by the Israel Science Foundation (ISF) and the Israel Cancer Research Foundation. L.J. is supported by the Edmond J. Safra Bioinformatics program at TAU. We thank UOB Tumor Cell Line Repository and W. Marston Linehan for providing us with the UOK262 cell lines, and A. King for editorial work.

**Author Contributions** C.F. and E.G. conceived the project, analysed the data and wrote the manuscript. C.F. and. E.D.M. generated the kidney cell lines. O.F., L.J., T.S. and E.R. built the metabolic network of the cells and predicted the synthetic lethal genes. J.A., I.P.M.T. and P.J.P. provided the $Fh1^{fl/fl}$ mice, generated the UOKpFH cell line and provided genotyping tools. G.K. and A.H. performed the bioinformatics and statistical analysis. L.Z. and D.G.W. performed the LC–MS analysis. K.N.R. and R.J.D. performed GC–MS and the BioProfile metabolic analyses. C.F. performed all other experiments with the help of B.C. M.M. performed the electron microscopy. All the authors discussed the results and commented on the manuscript.

**Author Information** Microarray data are deposited at the website http://bioinformatics.picr.man.ac.uk/vice/Welcome.vice, accession reference GE_EG(2). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to E.G. (e.gottlieb@beatson.gla.ac.uk).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Cell culture: *Fh1*$^{fl/fl}$ and *Fh1*$^{-/-}$ cell lines.** Epithelial kidney cells were obtained from *Fh1*$^{fl/fl}$ mice following the method described previously[19]. $2 \times 10^5$ cells were either left untreated or infected with Cre recombinase-encoding adenovirus (Ad5CMVCRE) using 300 p.f.u. per cell. After 4 days of infection, 100 cells were plated and grown onto 15-cm dishes. Individual colonies were harvested using paper cloning discs (Sigma) and transferred onto 24-well plates for clonal expansion. Cells were grown using DMEM (21969-035, Invitrogen) supplemented with 10% FBS, 2 mmol l$^{-1}$ glutamine, 1 mmol l$^{-1}$ pyruvate and 50 μg l$^{-1}$ uridine. UOK262 were cultured in DMEM (21969-035, Invitrogen) supplemented with 10% FBS, 2 mmol l$^{-1}$ glutamine, 1 mmol l$^{-1}$ pyruvate.

**UOK262 and UOKpFH cell lines.** UOK262 cells were grown using high glucose DMEM (Invitrogen) supplemented with 2 mM glutamine and 10% FBS. Full-length *FH* human cDNA was cloned into pEF1-V5-HIS. Transfected cells were selected in 1 mg ml$^{-1}$ G418 and single clones were expanded. Expression and subcellular localization of FH was confirmed by immunofluorescence using anti V5 antibody with a FITC-conjugated secondary antibody; mitochondria were stained using Mitotraker red and nuclei were stained with DAPI.

**Genomic DNA, mitochondrial DNA, RNA isolation and qPCR analyses.** $2 \times 10^5$ cells were harvested in RLT buffer (Qiagen) using a cell lifter and lysates were passed through QiaShredder columns (Qiagen). Genomic DNA and mRNA were then separated using the AllPrep micro kit (Qiagen) following manufacturer's instructions. Quantification of RNA was obtained on the Eppendorf Biophotometer using Eppendorf single-sealed cuvettes, UVette (Eppendorf UK Limited).

Genomic DNA was assessed to determine the different genotypes of the different cell lines using the primers F1, F3 and R1 described above. The PCR of the genomic DNA was performed using GoTaq DNA polymerase (Promega) and 0.5 μM primers following the manufacturer's instructions. The PCR program was: 15 min at 95 °C followed by 40 cycles of 45 s at 95 °C, 30 s at 55 °C and then 30 s at 72 °C. The PCR products were separated in a 2% agarose gel stained with ethidium bromide.

Mitochondrial DNA content was analysed from the genomic DNA extraction by comparing the expression levels of a nuclear-encoded single-copy gene, amyloid precursor protein (APP), to the DNA levels of the mtDNA encoded cytochrome *c* oxidase subunit II (COXII), as performed previously[20]. The following primers were used: *APP*: FW 5′-GGCTGCTGGCAGAACCCCAG-3′, RV 5′-CCCTGACGG GTCTGACTCCCA-3′; *COXII* FW 5′-AACTGATGCCATCCCAGGCCGA-3′, RV 5′-GGTGAAGGTGCCAGTGGGAATGT-3′.

For PCR analyses, 1 μg of mRNA was retro-transcribed into cDNA using M-Mulv Reverse Transcriptase and random hexamers and qPCR reactions were performed using a Sybr-green master mix as per the manufacturer's instructions (Dynamo qPCR kit, Finnzymes, Thermoscientific). In brief, 0.5 μM primers, ROX dye, and a 1 μl of a 1:10 dilution of cDNA in a final volume of 20 μl were used. The primers were designed by the Roche Universal Library (UPL, Roche, http://www.roche-applied-science.com/sis/rtpcr/upl/ezhome.html) and the sequences are indicated in Supplementary Table 5. Real-time PCR was performed on the 7500 Fast Real-Time PCR System (Life Technologies Corporation) and expression levels of the indicated genes were calculated using the $\Delta\Delta C_t$ method by the appropriate function of the software using actin as calibrator. The results were further validated using Polg2R as calibrator (data not shown). The PCR programme was: 15 min at 95 °C followed by 40 cycles of 10 s at 94 °C, 30 s at 55 °C and then 30 s at 72 °C. A final extension at 72 °C for 10 min was allowed before the melting curve, which was used to confirm the presence of single PCR products. For the semi-quantitative end-point RT–PCR the cDNA prepared as indicated above was cycled for 25 times and the PCR products were separated using a 2% agarose gel stained with ethidium bromide.

**Exon array analysis.** RNA obtained as described above was subjected to a Microarray analysis using a GeneChip Mouse Exon 1.0 ST Array (Affimetrix). The .Cel files were normalized and analysed in Partek Genomics Suite Software, version 6.5beta Copyright 2010. Interrogating and control core probe sets were imported and default RMA setting used, including quantile normalization, $\log_2$ transformation and mean probe-set summarization. In gene-level analysis the multiway analysis of variance (ANOVA) was used to identify significantly regulated genes from one of the experimental groups and linear contrasts performed between all pairs of experimental conditions. Multiple test corrections were performed for all calculated *P*-values. Step-up *P*-value, a *P*-value that controls the false discovery rate, and the fold change values were used for further selection of significantly differentially expressed genes.

**Microarray bioinformatics analysis.** Affymetrix Human Genome U133A .Cel files of 15 wild-type and 7 mutant fumarate hydratase samples were normalized and analysed in Partek Genomics Suite Software, version 6.5 Copyright 2010. RMA normalization and $\log_2$ transformation of the data was followed by *t*-test and multiple test corrections of *P*-values.

**Immunoblotting.** $2 \times 10^5$ cells were lysed in RIPA buffer (150 mM sodium chloride, 1.0% NP-40, 0.5% sodium deoxycholate, and 50 mM Tris, pH 8.0) supplemented with a 1:100 dilution of the protein inhibitor cocktail (Sigma). Protein concentration was determined using the Bicinchoninic Acid Assay (Thermoscientific) using BSA as standard. Equal amounts of protein were loaded into 4–12% Tris-Mops pre-casted gels (Invitrogen) and electrophoretically separated using MOPS running buffer (Invitrogen). After SDS–PAGE, proteins were blotted onto 0.22 μm nitrocellulose (Millipore) and probed with the following antibodies, all at 1:1,000 dilution in 5% non-fat milk: Fh1 (polyclonal rabbit AutogenBioclear NE054 /7S), Hmox1 (polyclonal rabbit SPA-896, Stressgen Enzo Life Science), actin (mouse monoclonal AC-40; Sigma). Membranes were then incubated with secondary HRP-conjugated antibodies at 1:1,000 dilution in 5% non-fat milk: anti-mouse (7076) and anti rabbit (7074), both from Cell Signaling Technology. ECL (RPN 2209, GE Healthcare UK) was used for chemiluminescence detection. For Hmox1 and actin western blot, a secondary donkey anti-rabbit (926 32213, Licor) and donkey anti-mouse (926 32212, Licor) IR800 at 1:1,000 dilution were used. The IR scanning was performed using the Licor Odyssey scanner (channel, 800; brightness, 50; contrast, 50; sensitivity, auto; resolution, 169.492 μm; pixel area, 0.02873 mm$^2$; intensity, 5) and acquired using Odyssey software version 3.

**Measurement of NADH and mitochondrial potential.** $2 \times 10^5$ cells were grown in 3.5-cm glass-bottom dishes (MatTek corporation) and loaded with 50 nmol l$^{-1}$ MitoTracker green (MTG; Invitrogen) and 10 nmol l$^{-1}$ tetramethyl rhodamine ethyl ester (TMRE) (Invitrogen) for 30 min at 37 °C. Cells were then left untreated or treated as indicated. The final concentrations and incubation time of the drugs are: 5 μmol l$^{-1}$ rotenone, 5 min; 10 μmol l$^{-1}$ carbonylcyanide-*m*-chlorophenyl hydrazone (CCCP) for 1 min. Images were collected with a Leica TCS SP2 AOBS inverted laser scanning microscope equipped with a Coherent 351–364 nm UV laser using a 63×/1.32.ph3 oil HCX PL APO objective, resolution 1,024 × 1,024 d.p.i., 8-bit image depth, 200 Hz scanning speed, and line average 2. The acquisition of the two-colour images was performed in a sequential scanning mode to minimize spectral bleed-through artefacts. Laser excitation was 364 nm for NADH autofluorescence, 488 nm for MTG and 568 nm for TMRE, whereas fluorescence emission was adjusted with AOBS at 390–486 nm for NADH, 501–546 nm for MTG and 580–620 nm for TMRE. To increase sensitivity of NADH detection, the pinhole aperture used for all channels was set at 2.62 airy units. The images were quantified for NADH levels using the Image Pro Plus software (Media Cybernetics) using MTG signal as a mask to outline the mitochondrial reticulum. Intensity of fluorescence of NADH determined in the outlined areas was then normalized to MTG fluorescence intensity. To better appreciate changes in NADH fluorescence intensity, a rainbow look-up table was applied to the relevant images using the appropriate function of ImageJ (ImageJ, US National Institutes of Health, http://rsb.info.nih.gov/ij/).

**Measurement of glucose, lactate and glutamine.** $2 \times 10^5$ cells were plated onto 6-cm dishes and cultured until 90% confluent. Each dish was then replenished with fresh medium and cultured for 16 h. After the culture period, the medium was recovered for determination of glucose, lactate and glutamine concentration using a BioProfile Basic4 Analyser (NOVA Biomedical) and compared to concentration of these metabolites in the starting medium. To correct for spontaneous glutamine degradation, an additional dish with no cells was cultured in parallel. Raw values of glutamine and glucose utilization, and lactate secretion, were converted to rates by correcting for the protein content at the end of the assay, and the duration of the assay in hours.

**Measurement of OCR and ECAR rate.** $2 \times 10^4$ *Fh1*$^{fl/fl}$ cells, $3 \times 10^4$ *Fh1*$^{-/-}$, or $3 \times 10^4$ UOK262 and UOK262pFH cells were plated onto XF24 plates (Seahorse Bioscience) and incubated at 37 °C, 5% $CO_2$ for 24 h. The medium was then replaced with 675 μl of unbuffered assay media (Seahorse Bioscience) supplemented with 2 mmol l$^{-1}$ glutamine, 25 mmol l$^{-1}$ glucose and 2% FBS (pH was adjusted to 7.4 using sodium hydroxide 0.5 M) and cells were then placed at 37 °C in a $CO_2$-free incubator. Basal oxygen consumption rate (OCR) and extracellular acidification rate (ECAR) were then recorded using the XF24 plate reader. At the end of the experiment 1 mmol l$^{-1}$ potassium cyanide or 1 μmol l$^{-1}$ antimycin A were added to measure mitochondria-independent oxygen consumption. Each cycle of measurement consisted of 3 min mixing, 3 min waiting and 4 min measuring. OCR and ECAR were normalized to cell number calculated at the end of the experiments. To obtain the mitochondrial-dependent OCR, only the cyanide-sensitive respiration was used. Homogeneous plating of the cells and cell count were assessed by fixing the cells with trichloroacetic acid 10% for 1 h at 4 °C and then staining the fixes cells with 0.47% solution of sulphorhodamine B (SRB) (Sigma).

**Determination of TCA cycle metabolites and isotopomer enrichment analysis by GC–MS.** Cells were cultured in DMEM supplemented with 10% FBS, 2 mmol

©2011 Macmillan Publishers Limited. All rights reserved

$l^{-1}$ sodium pyruvate, 50 μg $ml^{-1}$ uridine, 100 units $ml^{-1}$ penicillin and 100 μg $ml^{-1}$ streptomycin in 6-cm tissue culture dishes. Metabolic experiments were performed when cells were approximately 90% confluent. Medium for $^{13}C$ glucose labelling experiments contained 10 mM D-[U-$^{13}C$]-glucose (Cambridge Isotope Labs) and 4 mmol $l^{-1}$ unlabelled glutamine. Medium for $^{13}C$-gluamine labelling experiments contained 10 mmol $l^{-1}$ unlabelled glucose and 4 mmol $l^{-1}$ L-[U-$^{13}C$]-glutamine (Isotec). Medium containing $^{13}C$-labelled nutrients did not contain sodium pyruvate. Each dish of cells was rinsed in phosphate-buffered saline, and replenished with the labelling medium. After 16 h, cells were rinsed with normal saline, and metabolites were extracted in 0.5 ml 50% MeOH pre-chilled to −20 °C. 100 nmol of sodium oxobutyrate was added as a standard to the extracted metabolites. Extracts were concentrated under blown air and derivatized using Tri-Sil trimethylsilane reagent (Thermo Scientific). One microlitre of each derivatized sample was injected onto an Agilent 6980N GC coupled to an Agilent 5973N mass selective detector. GC–MS data were analysed using MSDChem Data Analysis Software (Agilent). Retention times of the sodium oxobutyrate standard, succinate, fumarate and malate were 11.4, 13.2, 14.2 and 18.7 min, respectively. Informative ions were monitored for derivatized, unenriched oxobutyrate (*m*/*z* 230), succinate (247), fumarate (245), and malate (335). Relative metabolite abundances were determined by normalizing abundances of each metabolite to the internal standard and to protein content. For mass isotopomer analysis, the abundance of each mass isotopomer (m+0, m+1, … m+*n*, where *n* is the number of carbons in the parent metabolite) was corrected for natural $^{13}C$ abundance using published methods[21].

**Measurement of bilirubin, ALA, lactate and TCA cycle intermediates by LC–MS.** $1 \times 10^6$ cells were plated onto 10-cm dishes and cultured in standard medium for 12 h. The medium was replaced by fresh standard medium or by medium without glucose or without glutamine for the indicated time. Fresh medium in the same dishes but without cells was incubated simultaneously and used as a reference. 100 μl of medium were diluted into 200 μl of acetonitrile (Fisher). The solution was centrifuged at 16,000*g* for 4 min at 4 °C. The supernatant was collected and transferred into autosampler vials for LC–MS determinations. Cells were lysed with a solution kept in dry ice/methanol (−80 °C) composed of 50% methanol and 30% acetonitrile in water and quickly scraped. The insoluble material was immediately pelleted in a cooled centrifuge (0 °C) and the supernatant was collected for subsequent LC-MS analysis. A ZIC-HILIC column (4.6 mm × 150 mm, guard column 2.1 mm × 20 mm, Merck) was used for LC separation using formic acid, water acetonitrile as component of the mobile phase.

**Measurement of $^{13}C$-labelled metabolites by LC–MS.** $1 \times 10^6$ cells were plated onto 10-cm dishes and cultured in standard medium for 12 h. The medium was replaced by fresh medium supplemented with 2 mM U-$^{13}C$-glutamine for the indicated time. Fresh medium in the same dishes but without cells was incubated simultaneously and used as a reference. Cells were then lysed with a solution kept in dry ice/methanol (−80 °C) composed of 50% methanol and 30% acetonitrile in water and quickly scraped. The insoluble material was immediately pelleted in a cooled centrifuge at 16,000*g* for 15 min at 0 °C and the supernatant was collected for subsequent LC–MS analysis. The amount of extraction solution was calculated according to the number of cells present in the sample dish, extrapolated using a 'counter dish' cultured in the same conditions of the sample dishes. A concentration of 1 ml per $2 \times 10^6$ cells was used in the extraction solutions. Intermediates were separated using a liquid chromatography system.

**Colony formation assay.** Clonogenic survival assay was performed as previously described[22] with minor modifications. The plating efficiency among different cell lines was different therefore 100 $Fh1^{fl/fl}$ cells and 600 $Fh1^{-/-}$ cells per each 10-cm dish were plated, whereas 150 UOK262 cells, 150 UOK262pFH cells and 600 UOK262 shHIFβ cells per each 6-well plate were plated. Twenty-four hours after plating, cells were treated with the indicated doses of ZnPP, 100 μmol $l^{-1}$ hemin (Sigma) or vehicle (DMSO for ZnPP or water for the other drug). UOK262 cells infected with shHIF1β plasmid were selected with 2 μg $ml^{-1}$ puromycin. Hemin was prepared by dissolving the drug in a solution of 0.2 M sodium carbonate, pH 10. The pH was then adjusted at 7.4 and the mixture was sterilized using 0.45 μm filters. The final concentration of the solution was determined using the molar extinction coefficient of hemin at $\lambda_{407\,nm}$ $\varepsilon = 36{,}000$ M $cm^{-1}$ (ref. 23). The medium was changed every 2 days for a total of 7 days treatment for murine cell lines and a total of 4 days of treatment for human cell lines. The medium was then replaced with fresh standard medium without drug or vehicle and cells were left to grow for 2–4 more days. The cells were then fixed as described[24] by adding ice-cold trichloroacetic acid onto the plate at the final concentration of 3% for 1 h at 4 °C. Cells were then washed with water and air-dried. Cells were then stained using 0.47% SRB for 30 min at room temperature. SRB was removed and cells were thoroughly washed with 1% acetic acid and air-dried. Plates were scanned at the resolution of 600 d.p.i., 24-bit saved as TIFF files and analysed using the object count function of the ImageJ software.

**ShRNA of *Hmox1*.** The lentiviral shRNA plasmids against *Hmox1* (NM_010442) and the non-targeted shGFP, and a scramble sequence were all purchased from Open Biosystems, TRC1 library (Thermoscientific). The three independent shRNAs were identified as follows: 1, TRCN0000071578; 2, TRCN0000071579; 3, TRCN0000071580 and the sequences are shown in Supplementary Table 6. The viral supernatant was obtained from the filtered growth media of the packaging cells HEK293T transfected with appropriate packaging plasmids and either shRNA for *Hmox*1 or control pLKO plasmids, using calcium phosphate procedure. $2 \times 10^5$ $Fh1^{fl/fl}$, $Fh1^{-/-}$ and UOK262 cells were then plated on 6-well plates and infected with the viral supernatant prepared as indicated above. After 2 days, the medium was replaced by selection medium containing 4 μg $ml^{-1}$ puromycin. Cells were then allowed to grow under selection for 2 days and then an equal number of cells was plated onto 6-well plates and grown using the selection media. After a week cells where fixed using SRB following the procedure described above.

**shRNA of *HIF1β* (ARNT).** $2 \times 10^5$ cells were infected with shRNA for *HIF1β* (NM_001668.2) and control plasmid using lentiviral infection. HEK293T and relevant packaging plasmids were used to produce the lentiviral suspension. shRNA plasmids were purchased from Openbiosystems and the clone number TRCN0000003819 (for the sequence see Supplementary Table 6) was used for subsequent infections. After infection, cells were selected using 4 μg $ml^{-1}$ puromycin.

**Statistical analyses.** The data (mean ± s.e.m.) are representative of 3–5 independent experiments, performed in technical triplicates if not differently indicated. Data were analysed and presented with Graphpad Prism 5.01 software (GraphPad Software Inc.).


20. Malena, A., Loro, E., Di Re, M., Holt, I. J. & Vergani, L. Inhibition of mitochondrial fission favours mutant over wild-type mitochondrial DNA. *Hum. Mol. Genet.* **18,** 3407–3416 (2009).
21. Fernandez, C. A., Des Rosiers, C., Previs, S. F., David, F. & Brunengraber, H. Correction of $^{13}C$ mass isotopomer distributions for natural stable isotope abundance. *J. Mass Spectrom.* **31,** 255–262 (1996).
22. Franken, N. A., Rodermond, H. M., Stap, J., Haveman, J. & van Bree, C. Clonogenic assay of cells *in vitro*. *Nature Protocols* **1,** 2315–2319 (2006).
23. Uc, A., Stokes, J. B. & Britigan, B. E. Heme transport exhibits polarity in Caco-2 cells: evidence for an active and membrane protein-mediated process. *Am. J. Physiol. Gastrointest. Liver Physiol.* **287,** G1150–G1157 (2004).
24. Vichai, V. & Kirtikara, K. Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nature Protocols* **1,** 1112–1116 (2006).


©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10373

# Broad neutralization coverage of HIV by multiple highly potent antibodies

Laura M. Walker[1]*, Michael Huber[1,2]*, Katie J. Doores[1,2]*, Emilia Falkowska[1,2], Robert Pejchal[3], Jean-Philippe Julien[3], Sheng-Kai Wang[4], Alejandra Ramos[1], Po-Ying Chan-Hui[5], Matthew Moyle[5], Jennifer L. Mitcham[5], Phillip W. Hammond[5], Ole A. Olsen[5], Pham Phung[6], Steven Fling[7], Chi-Huey Wong[4], Sanjay Phogat[7], Terri Wrin[6], Melissa D. Simek[7], Protocol G. Principal Investigators†, Wayne C. Koff[7], Ian A. Wilson[3], Dennis R. Burton[1,2] & Pascal Poignard[1,7]

**Broadly neutralizing antibodies against highly variable viral pathogens are much sought after to treat or protect against global circulating viruses. Here we probed the neutralizing antibody repertoires of four human immunodeficiency virus (HIV)-infected donors with remarkably broad and potent neutralizing responses and rescued 17 new monoclonal antibodies that neutralize broadly across clades. Many of the new monoclonal antibodies are almost tenfold more potent than the recently described PG9, PG16 and VRC01 broadly neutralizing monoclonal antibodies and 100-fold more potent than the original prototype HIV broadly neutralizing monoclonal antibodies[1–3]. The monoclonal antibodies largely recapitulate the neutralization breadth found in the corresponding donor serum and many recognize novel epitopes on envelope (Env) glycoprotein gp120, illuminating new targets for vaccine design. Analysis of neutralization by the full complement of anti-HIV broadly neutralizing monoclonal antibodies now available reveals that certain combinations of antibodies should offer markedly more favourable coverage of the enormous diversity of global circulating viruses than others and these combinations might be sought in active or passive immunization regimes. Overall, the isolation of multiple HIV broadly neutralizing monoclonal antibodies from several donors that, in aggregate, provide broad coverage at low concentrations is a highly positive indicator for the eventual design of an effective antibody-based HIV vaccine.**

Most successful antiviral vaccines elicit neutralizing antibodies as a correlate of protection[4,5]. For highly variable viruses—such as HIV, hepatitis C virus (HCV) and, to a lesser extent, influenza—vaccine design efforts have been hampered by the difficulties associated with eliciting neutralizing antibodies that are effective against the enormous diversity of global circulating isolates (that is, broadly neutralizing antibodies)[6,7]. However, for HIV for example, 10–30% of infected individuals do, in fact, develop broadly neutralizing sera and protective broadly neutralizing monoclonal antibodies have been isolated from infected donors[1,2,8–12]. It has been suggested that, given the appropriate immunogen, it should be possible to elicit these types of responses by vaccination[13] and understanding the properties of broadly neutralizing monoclonal antibodies has become a major goal in research on highly variable viruses.

We have previously screened sera from approximately 1,800 HIV-infected donors for neutralization breadth and potency, designating the top 1% as 'elite neutralizers', based on a score incorporating both breadth and potency[14]. In this study, we set out to isolate broadly neutralizing monoclonal antibodies from the top four elite neutralizers (Supplementary Table 1) by screening antibody-containing memory B cell supernatants for broad neutralizing activity using a recently described high-throughput functional approach[2]. Antibody variable genes were rescued from B-cell cultures that showed cross-clade neutralizing activity and expressed as full-length IgGs. Analysis of the sequences revealed that all of the monoclonal antibodies isolated from each individual donor belong to a distant, but clonally related cluster of antibodies (Supplementary Table 2). Because it has been proposed that antibodies from HIV-infected patients are often polyreactive[15,16], we tested the new monoclonal antibodies for binding to a panel of antigens and showed that they were not polyreactive (Supplementary Fig. 2).

The potency and breadth of the monoclonal antibodies were next assessed on a 162-pseudovirus panel representing all major circulating HIV subtypes (Fig. 1 and Supplementary Tables 3 and 4)[2]. All of the monoclonal antibodies exhibited cross-clade neutralizing activity, but more strikingly, several showed exceptional potency. The median antibody concentrations required to inhibit HIV activity by 50% or 90% ($IC_{50}$ and $IC_{90}$ values), of PGT monoclonal antibodies 121–123 and 125–128 were almost tenfold lower (that is, more potent) than the recently described PG9, PG16, VRC01 and PGV04 broadly neutralizing monoclonal antibodies[1,2] (E. Falkowska *et al.*, manuscript in preparation), and approximately 100-fold lower than other broadly neutralizing monoclonal antibodies described earlier (Fig. 1). At

**a**

| | Median $IC_{50}$ |
|---|---|
| PGT121 | 0.03 |
| PGT122 | 0.05 |
| PGT123 | 0.03 |
| PGT125 | 0.04 |
| PGT126 | 0.04 |
| PGT127 | 0.08 |
| PGT128 | 0.02 |
| PGT130 | 0.16 |
| PGT131 | 0.52 |
| PGT135 | 0.17 |
| PGT136 | 7.81 |
| PGT137 | 3.46 |
| PGT141 | 0.35 |
| PGT142 | 0.21 |
| PGT143 | 0.31 |
| PGT144 | 2.06 |
| PGT145 | 0.29 |
| PG9 | 0.23 |
| VRC01 | 0.32 |
| PGV04 | 0.20 |
| b12 | 2.82 |
| 2G12 | 2.38 |
| 4E10 | 3.41 |

Key: < 0.2 (μg ml$^{-1}$); 0.2–2 (μg ml$^{-1}$); 2–20 (μg ml$^{-1}$)

**b**

| | Viruses neutralized | | |
|---|---|---|---|
| | $IC_{50}$ < 50 μg ml$^{-1}$ | $IC_{50}$ < 1 μg ml$^{-1}$ | $IC_{50}$ < 0.1 μg ml$^{-1}$ |
| PGT121 | 70 | 57 | 44 |
| PGT122 | 65 | 48 | 36 |
| PGT123 | 67 | 54 | 40 |
| PGT125 | 52 | 40 | 32 |
| PGT126 | 60 | 50 | 40 |
| PGT127 | 50 | 37 | 27 |
| PGT128 | 72 | 60 | 50 |
| PGT130 | 52 | 35 | 23 |
| PGT131 | 40 | 23 | 13 |
| PGT135 | 33 | 23 | 13 |
| PGT136 | 16 | 6 | 3 |
| PGT137 | 22 | 8 | 4 |
| PGT141 | 56 | 36 | 15 |
| PGT142 | 57 | 40 | 23 |
| PGT143 | 56 | 37 | 17 |
| PGT144 | 38 | 16 | 3 |
| PGT145 | 78 | 52 | 27 |
| PG9 | 77 | 54 | 29 |
| VRC01 | 93 | 74 | 20 |
| PGV04 | 88 | 65 | 25 |
| b12 | 34 | 10 | 2 |
| 2G12 | 32 | 11 | 1 |
| 4E10 | 96 | 13 | 1 |

Key: > 90 %; 60 - 90 %; 30 - 60 %; 1 - 30 %

**Figure 1 | Neutralization activity of the newly identified PGT antibodies.** **a**, Median neutralization potency against viruses neutralized with an $IC_{50} < 50$ μg ml$^{-1}$. **b**, Neutralization breadth at different $IC_{50}$ cut-offs.

[1]Department of Immunology and Microbial Science and IAVI Neutralizing Antibody Center, The Scripps Research Institute, La Jolla, California 92037, USA. [2]Ragon Institute of MGH, MIT, and Harvard, Cambridge, Massachusetts 02129, USA. [3]Department of Molecular Biology, Skaggs Institute for Chemical Biology, and IAVI Neutralizing Antibody Center, The Scripps Research Institute, La Jolla, California 92037, USA. [4]Department of Chemistry, The Scripps Research Institute, La Jolla, California 92037, USA. [5]Theraclone Sciences, Inc., Seattle, Washington 98104, USA. [6]Monogram Biosciences, Inc., South San Francisco, California 94080, USA. [7]International AIDS Vaccine Initiative, New York, New York 10038, USA.

*These authors contributed equally to this work.

†Lists of authors and affiliations appear in Supplementary Information.



©2011 Macmillan Publishers Limited. All rights reserved

concentrations less than 0.1 μg ml$^{-1}$, these monoclonal antibodies still neutralized 27% to 50% of viruses in the panel (Fig. 1). Although PGT monoclonal antibodies 135, 136 and 137 showed a lesser neutralization breadth than the other monoclonal antibodies, they all still potently neutralized over 30% of the clade C viruses on the panel (Supplementary Fig. 2 and Supplementary Table 3b). This result is significant considering that HIV clade C predominates in sub-Saharan Africa and accounts for more than 50% of all HIV infections worldwide.

Interestingly, many of the clonally related monoclonal antibodies exhibited differing degrees of overall neutralization potency. For example, the median $IC_{50}$ values of PGT monoclonal antibodies 131, 136, 137 and 144 were approximately 10- to 50-fold higher than those of their somatically related sister clones (Fig. 1). Also, in some cases, the somatically related monoclonal antibodies exhibited similar neutralization potency but differing degrees of neutralization breadth against the panel of viruses tested (Fig. 1 and Supplementary Tables 3 and 4). For example, PGT 128 neutralized with comparable overall potency but significantly greater neutralization breadth than the clonally related PGT 125, 126 and 127 monoclonal antibodies (Fig. 1 and Supplementary Tables 3 and 4). Overall, these observations suggest that serum neutralization breadth may develop from the successive selection of somatic variants that bind to a modified epitope or a slightly different Env conformation expressed on virus escape variants. Comparison of the neutralization profiles of the monoclonal antibodies isolated from a given donor with that from the corresponding serum revealed that the isolated monoclonal antibodies could largely recapitulate the serum neutralization breadth and potency (Fig. 2 and Supplementary Fig. 3).

We next sought to gain information on the epitopes recognized by the newly isolated broadly neutralizing monoclonal antibodies. Enzyme-linked immunosorbent (ELISA) assays indicated that PGT monoclonal antibodies 121–123, 125–128, 130, 131 and 135–137 bound to monomeric gp120 (Supplementary Table 5). In contrast, the PGT 141–145 broadly neutralizing monoclonal antibodies exhibited a strong preference for membrane-bound, trimeric HIV Env (Supplementary Fig. 4). On the basis of this result, we postulated that these broadly neutralizing monoclonal antibodies bound to quaternary epitopes similar to those of the recently described PG9 and PG16 broadly neutralizing monoclonal antibodies[2]. Indeed, this hypothesis was confirmed by competition studies, N160K sensitivity and, for PGT monoclonal antibodies 141–144, an inability to neutralize JR-CSF pseudoviruses expressing homogenous $Man_9GlcNAc_2$ glycans[17] (Supplementary Fig. 5).

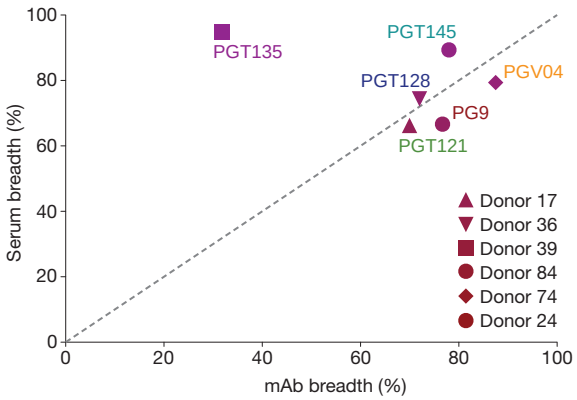
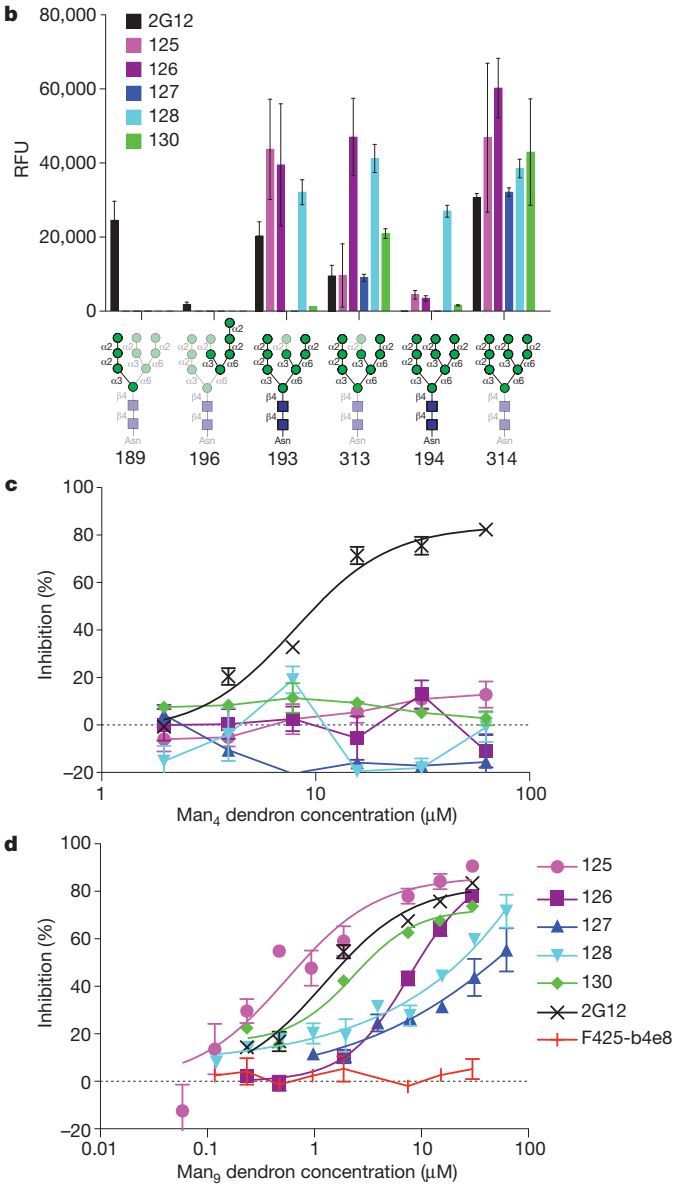



**Figure 2 | Key monoclonal antibodies fully recapitulate serum neutralization by the corresponding donor serum.** Serum breadth was correlated with the breadth of the broadest monoclonal antibody (mAb) for each donor (percentage of viruses neutralized at 50% neutralizing titre ($NT_{50}$) > 100 or $IC_{50}$ < 50 μg ml$^{-1}$, respectively). Of note, monoclonal antibodies isolated from donor 39 could not completely recapitulate the serum neutralization breadth.

To define the epitopes recognized by the remaining PGT antibodies, competition ELISA assays were carried out with a panel of well-characterized neutralizing and non-neutralizing antibodies (Fig. 3a). Unexpectedly, all of the remaining antibodies (PGT monoclonal antibodies 121–123, 125–128, 130, 131 and 135–137) competed with the glycan-specific broadly neutralizing monoclonal antibody 2G12. This result was surprising given that 2G12 had previously formed its own unique competition group. All of the monoclonal antibodies, except for PGT monoclonal antibodies 135, 136 and 137, also

**a**

| | Competitor antibody | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| mAb | sCD4 | b12 | 2G12 | F425 | X5 | PG9 | 121 | 125 | 130 | 135 |
| 121 | + | − | +++ | ++ | + | +++ | +++ | +++ | + | ++ |
| 122 | − | − | ++ | + | − | +++ | +++ | ++ | + | +++ |
| 123 | + | − | +++ | +++ | − | +++ | +++ | +++ | ++ | +++ |
| 125 | ++ | − | +++ | ++ | − | + | +++ | +++ | + | + |
| 126 | + | − | +++ | + | − | + | +++ | ++ | + | +++ |
| 127 | + | − | +++ | ++ | − | + | +++ | +++ | + | +++ |
| 128 | − | − | +++ | + | − | + | ++ | + | + | ++ |
| 130 | ++ | − | ++ | +++ | − | + | + | +++ | +++ | − |
| 131 | +++ | − | +++ | +++ | + | + | +++ | +++ | +++ | + |
| 135 | − | ++ | +++ | − | − | − | ++ | ++ | − | +++ |
| 136 | − | ++ | +++ | − | − | − | ++ | ++ | − | +++ |
| 137 | − | + | +++ | − | − | − | − | − | − | +++ |






©2011 Macmillan Publishers Limited. All rights reserved

competed with a V3-loop-specific monoclonal antibody and failed to bind to gp120ΔV3, suggesting that their epitopes were in proximity to or contiguous with V3 (Fig. 3a and Supplementary Table 5). Deglycosylation of gp120 with Endo H abolished binding by all the monoclonal antibodies, indicating that certain oligomannose glycans were important for epitope recognition (Supplementary Table 5). Competition of these monoclonal antibodies with 2G12 and lack of binding to deglycosylated gp120 prompted us to investigate whether these antibodies contacted glycans directly. Glycan array analysis revealed that PGT monoclonal antibodies 125–128 and 130 bound specifically to both $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$, whereas the remaining antibodies showed no detectable binding to high-mannose glycans (Fig. 3b). Interestingly, binding of PGT monoclonal antibodies 125–128 and 130 to gp120 was competed by $Man_9$ but, unlike 2G12, was not competed by monomeric mannose or $Man_4$ (D1 arm of $Man_9GlcNAc_2$) (Fig. 3c, d), suggesting a different mode of glycan recognition. Furthermore, in contrast to 2G12, no evidence was found for domain exchange and monomeric Fab fragments still exhibited potent neutralizing activity (Supplementary Fig. 7 and data not shown).

To define further the epitopes recognized by the monoclonal antibodies, neutralizing activity against a large panel of HIV-$1_{JR\text{-}CSF}$ variants incorporating single alanine substitutions was assessed using a single round of replication pseudovirus assay (Supplementary Table 6). In the panel of mutants, the N-linked glycans at positions 332 and/or 301 were important for neutralization by PGT monoclonal antibodies 125–128, 130 and 131, suggesting their direct involvement in epitope formation. The apparent dependency on so few glycans indicates that, although these PGT monoclonal antibodies contact $Man_{8-9}GlcNAc_2$ glycans directly, their arrangement in the context of gp120 is critical for high-affinity glycan recognition and neutralization potency. This is further highlighted by the inability of the PGT monoclonal antibodies to neutralize simian immunodeficiency virus (SIV) strain $SIV_{MAC239}$, HIV-2 or HCV, which show a high level of glycosylation (data not shown). Interestingly, although PGT monoclonal antibodies 121–123 failed to exhibit detectable binding to high-mannose glycans and be competed by mannose sugars (Supplementary Fig. 6), the only substitutions that completely abolished neutralization by these monoclonal antibodies were those that resulted in removal of the glycan at position 332. Although structural studies will be required to fully define the epitopes recognized by these antibodies, the above results indicate either that the PGT monoclonal antibodies 121–123 bind to a protein epitope along the gp120 polypeptide backbone that is conformationally dependent on the N332 glycan or that the glycan contributes more strongly to binding in the context of the intact protein.

Vaccines against pathogens with low antigenic diversity, such as hepatitis B virus or measles, commonly achieve 90–95% efficacy[18]. Similarly, the influenza vaccine achieves 85–90% efficacy in years when the vaccine and circulating seasonal strain are well matched[19,20]. However, efficacy drops severely in years when there is a mismatch between the vaccine and circulating strain. In the case of HIV, the global diversity of circulating viruses is such that the match between the prophylactic antibodies and the circulating viruses—that is, the antibody viral coverage—will be crucial for the degree of efficacy of active or passive prophylaxis approaches. As yet, although the recent

**Figure 3 | Epitope mapping of PGT antibodies.** **a**, Competition of PGT monoclonal antibodies with sCD4 (soluble CD4), b12 (anti-CD4 binding site), 2G12 (anti-glycan), F425/b4e8 (anti-V3), X5 (CD4-induced), PG9 (anti-V1/V2 and V3, quaternary) and each other. Competition assays were performed by ELISA using $gp120_{Bal}$ or $gp120_{JR\text{-}FL}$, except for the PG9 competition assay, which was performed on the surface of $JR\text{-}FL_{E168K}$ or JR-CSF transfected cells. Boxes are coded as follows: +++, 75–100% competition; ++, 50–75% competition; +, 25–50% competition; −, <25% competition. Experiments were performed in duplicate, and data represent an average of at least two independent experiments. **b**, Glycan microarray analysis (Consortium for Functional Glycomics (CFG), version 5.0) reveals that PGT monoclonal antibodies 125, 126, 127, 128 and 130 contact $Man_8$ (313), $Man_8GlcNAc_2$ (193), $Man_9$ (314) and $Man_9GlcNAc_2$ (194) glycans directly. Only glycan structures with RFU (relative fluorescent units) > 3,000 are shown. PGT-131 showed no detectable binding to the CFG glycan array but bound to $Man_9$-oligodendrons[30] (data not shown). Error bars represent standard deviation. **c**, **d**, Binding of PGT monoclonal antibodies 125, 126, 127, 128 and 130 to gp120 is competed by $Man_9$ oligodendrons but not $Man_4$ oligodendrons. Binding of 131 to immobilized gp120 was too low to measure any competition. Error bars represent standard error of the mean.

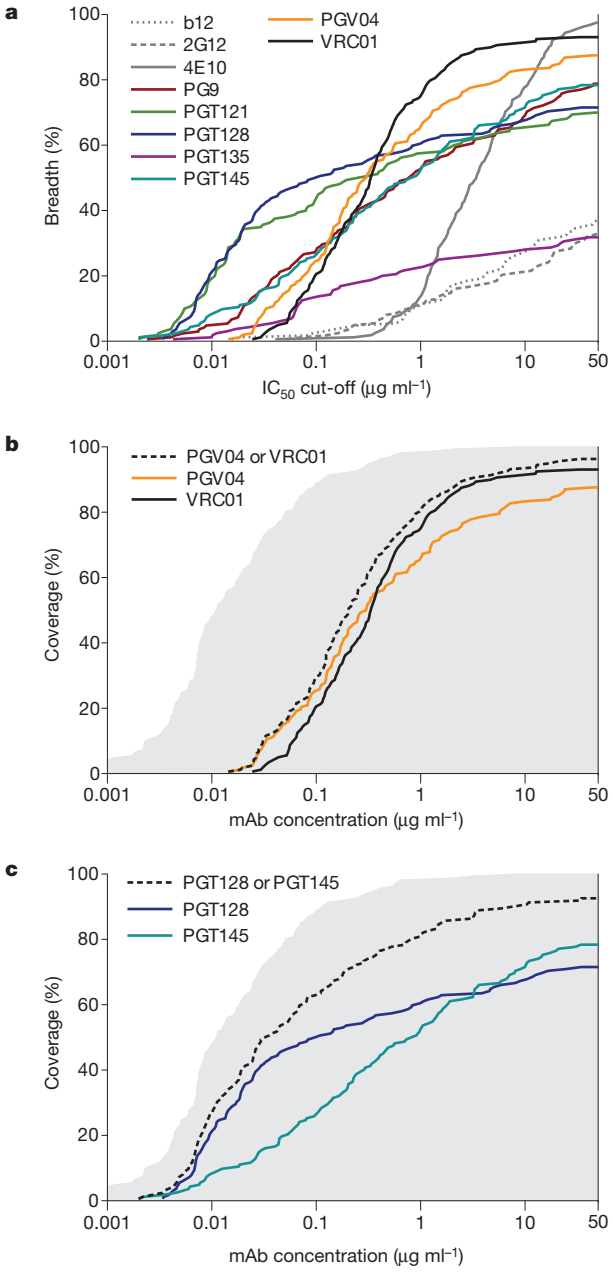



**Figure 4 | Certain antibodies or antibody combinations are able to cover a broad range of HIV isolates at low, vaccine-achievable concentrations.** **a**, Cumulative frequency distribution of $IC_{50}$ values of broadly neutralizing monoclonal antibodies tested against a 162-virus panel. The *y*-axis shows the cumulative frequency of $IC_{50}$ values up to the concentration shown on the *x*-axis and can therefore also be interpreted as the breadth at a specific $IC_{50}$ cut-off. **b**, **c**, Percentage of viruses covered by single monoclonal antibodies (solid lines) or by at least one of the monoclonal antibodies in dual combinations of breadth (dashed black lines) dependent on individual concentrations. The grey area in both panels is the coverage of 26 monoclonal antibodies tested on the 162-virus panel (PGT121–123, PGT125–128, PGT130–131, PGT135–137, PGT141–145, PG9, PG16, PGC14, VRC01, PGV04, b12, 2G12, 4E10, 2F5) and depicts the theoretical maximal achievable coverage known to date.



©2011 Macmillan Publishers Limited. All rights reserved

RV144 trial has led to speculation that some degree of protection against HIV may be achieved through extra-neutralizing activities of antibodies, such as antibody-dependent cell-mediated cytotoxicity or phagocytosis, the strongest evidence for protection is for neutralizing antibodies in non-human primate models using simian-human immunodeficiency virus (SHIV) challenge[21–25]. Passive administration of neutralizing antibodies in these models suggests that a serum antibody concentration of approximately or greater than 100 times the *in vitro* pseudovirus assay $IC_{50}$ is required to achieve a meaningful level of protection[21–25]. Therefore, if a vaccine elicits a serum broadly neutralizing antibody concentration on the order of 10 $\mu g\,ml^{-1}$ (ref. 26) and if an $IC_{50}$:protective-serum-concentration ratio of 1:100 is assumed, then protection would be only achieved against viruses for which the broadly neutralizing antibody $IC_{50}$ is lower than 0.1 $\mu g\,ml^{-1}$. As a second more conservative scenario, for an $IC_{50}$:protective-serum-concentration ratio of 1:500, protection would be achieved against viruses for which the broadly neutralizing antibody $IC_{50}$ is lower than 0.02 $\mu g\,ml^{-1}$. As shown in Fig. 4, although various broadly neutralizing monoclonal antibodies show breadth at high concentrations, viral coverage often drops sharply at lower concentrations. Therefore, if elicited or delivered singly, only the most potent antibodies, such as 121 and 128, would be able to achieve a meaningful level of viral coverage, in particular at concentrations corresponding to the more conservative scenario given above. As broadly neutralizing monoclonal antibodies show different and in some cases complementary breadth, we further looked at the theoretical coverage achieved by antibody combinations. For the two $IC_{50}$:protective-serum-concentration ratios above, a combination of PGV04 and VRC01, the two most potent CD4 binding site broadly neutralizing monoclonal antibodies, would provide protection against 29% and 2% of viruses, respectively (Fig. 4b). In contrast, for a vaccine eliciting antibodies with high potency and favourable non-overlapping breadth, such as 128 and 145, coverage would be achieved against 63% and 40% of viruses for the two scenarios (Fig. 4c). Several combinations of two broadly neutralizing monoclonal antibodies, including those directed to overlapping epitopes, can yield this degree of coverage (Supplementary Fig. 8). In addition, a combination of all of the broadly neutralizing monoclonal antibodies would cover 89% and 62% of viruses, correspondingly. Coverage against such a large proportion of viruses would probably have an important impact on the pandemic.

An effective vaccine against HIV will probably require the elicitation of a combination of complementary potent neutralizing antibodies. The demonstration that large numbers of potent and diverse broadly neutralizing monoclonal antibodies can be isolated from several different individuals provides grounds for renewed optimism that an antibody-based vaccine may be achievable.

## METHODS SUMMARY

Activated memory B-cell supernatants were screened in a high-throughput format for neutralization activity using a micro-neutralization assay, as described[2]. Heavy- and light-chain variable regions were isolated from B-cell lysates of selected neutralizing hits by reverse transcription from RNA followed by multiplex PCR amplification using family-specific V-gene primer sets. For some antibodies, traditional cloning methods were used for antibody isolation, as described[2]. For other antibodies, amplicons from each lysate were uniquely tagged with multiplex identifier (MID) sequences and 454 sequencing regions (Roche). Single rounds of replication pseudovirus neutralization assays and cell surface binding assays were performed as described previously[2,27,28]. Glycan reactivities were profiled on a printed glycan microarray (version 5.0 from the Consortium for Functional Glycomics) as described previously[29].

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 20 April; accepted 14 July 2011.**
**Published online 17 August 2011.**

1. Wu, X. *et al.* Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329,** 856–861 (2010).
2. Walker, L. M. *et al.* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326,** 285–289 (2009).
3. Binley, J. M. *et al.* Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J. Virol.* **78,** 13232–13252 (2004).
4. Amanna, I. J., Messaoudi, I. & Slifka, M. K. Protective immunity following vaccination: how is it defined? *Hum. Vaccin.* **4,** 316–319 (2008).
5. Plotkin, S. A. Immunologic correlates of protection induced by vaccination. *Pediatr. Infect. Dis. J.* **20,** 63–75 (2001).
6. Barouch, D. H. Challenges in the development of an HIV-1 vaccine. *Nature* **455,** 613–619 (2008).
7. Karlsson Hedestam, G. B. *et al.* The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. *Nature Rev. Microbiol.* **6,** 143–155 (2008).
8. Stamatatos, L., Morris, L., Burton, D. R. & Mascola, J. R. Neutralizing antibodies generated during natural HIV-1 infection: good news for an HIV-1 vaccine? *Nature Med.* **15,** 866–870 (2009).
9. Trkola, A. *et al.* Cross-clade neutralization of primary isolates of human immunodeficiency virus type 1 by human monoclonal antibodies and tetrameric CD4-IgG. *J. Virol.* **69,** 6609–6617 (1995).
10. Stiegler, G. *et al.* A potent cross-clade neutralizing human monoclonal antibody against a novel epitope on gp41 of human immunodeficiency virus type 1. *AIDS Res. Hum. Retroviruses* **17,** 1757–1765 (2001).
11. Burton, D. R. *et al.* Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody. *Science* **266,** 1024–1027 (1994).
12. Kwong, P. D. & Wilson, I. A. HIV-1 and influenza antibodies: seeing antigens in new ways. *Nature Immunol.* **10,** 573–578 (2009).
13. Schief, W. R., Ban, Y. E. & Stamatatos, L. Challenges for structure-based HIV vaccine design. *Curr. Opin. HIV AIDS* **4,** 431–440 (2009).
14. Simek, M. D. *et al.* Human immunodeficiency virus type 1 elite neutralizers: individuals with broad and potent neutralizing activity identified by using a high-throughput neutralization assay together with an analytical selection algorithm. *J. Virol.* **83,** 7337–7348 (2009).
15. Haynes, B. F. *et al.* Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. *Science* **308,** 1906–1908 (2005).
16. Mouquet, H. *et al.* Polyreactivity increases the apparent affinity of anti-HIV antibodies by heteroligation. *Nature* **467,** 591–595 (2010).
17. Walker, L. M. *et al.* A limited number of antibody specificities mediate broad and potent serum neutralization in selected HIV-1 infected individuals. *PLoS Pathog.* **6,** e1001028 (2010).
18. Plotkin, S. A. Orenstein, W. A. & Offit, P. A. *Vaccines* (Elsevier Health Sciences, Philadelphia, 2008).
19. Bridges, C. B. *et al.* Effectiveness and cost-benefit of influenza vaccination of healthy working adults: a randomized controlled trial. *J. Am. Med. Assoc.* **284,** 1655–1663 (2000).
20. Herrera, G. A. *et al.* Influenza vaccine effectiveness among 50–64-year-old persons during a season of poor antigenic match between vaccine and circulating influenza virus strains: Colorado, United States, 2003–2004. *Vaccine* **25,** 154–160 (2007).
21. Parren, P. W. *et al.* Antibody protects macaques against vaginal challenge with a pathogenic R5 simian/human immunodeficiency virus at serum levels giving complete neutralization *in vitro*. *J. Virol.* **75,** 8340–8347 (2001).
22. Nishimura, Y. *et al.* Determination of a statistically valid neutralization titre in plasma that confers protection against simian-human immunodeficiency virus challenge following passive transfer of high-titered neutralizing antibodies. *J. Virol.* **76,** 2123–2130 (2002).
23. Hessell, A. J. *et al.* Effective, low-titer antibody protection against low-dose repeated mucosal SHIV challenge in macaques. *Nature Med.* **15,** 951–954 (2009).
24. Hessell, A. J. *et al.* Broadly neutralizing human anti-HIV antibody 2G12 is effective in protection against mucosal SHIV challenge even at low serum neutralizing titers. *PLoS Pathog.* **5,** e1000433 (2009).
25. Willey, R., Nason, M. C., Nishimura, Y., Follmann, D. A. & Martin, M. A. Neutralizing antibody titers conferring protection to macaques from a simian/human immunodeficiency virus challenge using the TZM-bl assay. *AIDS Res. Hum. Retroviruses* **26,** 89–98 (2010).
26. Martell, B. A. *et al.* Cocaine vaccine for the treatment of cocaine dependence in methadone-maintained patients: a randomized, double-blind, placebo-controlled efficacy trial. *Arch. Gen. Psychiatry* **66,** 1116–1123 (2009).
27. Pantophlet, R. *et al.* Fine mapping of the interaction of neutralizing and nonneutralizing monoclonal antibodies with the CD4 binding site of human immunodeficiency virus type 1 gp120. *J. Virol.* **77,** 642–658 (2003).
28. Li, M. *et al.* Human immunodeficiency virus type 1 *env* clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *J. Virol.* **79,** 10108–10125 (2005).
29. Blixt, O. *et al.* Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc. Natl Acad. Sci. USA* **101,** 17033–17038 (2004).
30. Wang, S. K. *et al.* Targeting the carbohydrates on HIV-1: interaction of oligomannose dendrons with human monoclonal antibody 2G12 and DC-SIGN. *Proc. Natl Acad. Sci. USA* **105,** 3690–3695 (2008).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank N. Schultz, K. Le and V. Thaney for technical assistance at The Scripps Research Institute. We also thank C. Corbaci, S. M. Eagol and C. Williams for assistance with Fig. 3. We would also like to thank the research staff members of the Theraclone I-STAR Discovery and Antibody Purification Team, and K. Limoli, S. Jauregui, K. Tran, P. Sitou, J. Perea, B. Huang and P. Guinto at Monogram



©2011 Macmillan Publishers Limited. All rights reserved

BioSciences. We thank R. McBride for assistance with the glycan array and are grateful to W. Olson for providing purified KNH1144 SOSIP trimer, R. Wyatt for providing the YU2 Foldon trimer and J. Mascola for providing monoclonal antibody VRC01. In addition, we would like to thank all of the International AIDS Vaccine Initiative (IAVI) Protocol G project, clinical and site team members and all of the Protocol G clinical investigators, specifically, G. Miiro, A. Pozniak, D. McPhee, O. Manigart, E. Karita, A. Inwoley, W. Jaoko, J. DeHovitz, L.-G. Bekker, P. Pitisuttithum, R. Paris, J. Serwanga and S. Allen. This work was supported by IAVI, NIAID, the Ragon Institute and the United States for International Development (USAID). The contents are the responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government.

**Author Contributions** Project planning was performed by L.M.W., M.H., K.J.D., E.F., P.-Y.C.-H., R.P., J-P.J., T.W., S.P., M.D.S., W.K., I.A.W., M.M., S.F., D.R.B. and P.P.; experimental work by L.M.W., M.H., K.J.D., E.F., R.P., J.-P. J., P.-Y.C.-H., S.-K.W., A.R., J.L.M., P.W.H., O.A.O. and T.W.; data analysis by M.H., L.M.W., K.J.D., E.F., T.W., P.-Y.C.-H., D.R.B. and P.P.; composition of the first draft of the manuscript by L.M.W. and D.R.B.; major manuscript revisions by M.H., K.J.D. and P.P.; manuscript layout and revisions by R.P., J.-P.J., E.F., W.C.K. and I.A.W.

**Author Information** Gene sequences of the reported antibodies and the primers used for antibody isolation have been deposited under GenBank accession numbers JN201894–JN201927. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.R.B. (burton@scripps.edu) or P.P. (poignard@scripps.edu).



©2011 Macmillan Publishers Limited. All rights reserved

## METHODS

**Antibodies and antigens.** The following antibodies and reagents were procured by the IAVI Neutralizing Antibody Consortium: antibody 2G12 (Polymun Scientific), antibody F425/b4E8 (provided by L. Cavacini, Beth Israel Deaconess Medical Center), soluble CD4 (Progenics), HxB2 gp120, SF162 gp120, BaL gp120, JR-FL gp120, JR-CSF gp120 and YU2 gp120 (provided by G. Stewart-Jones, Oxford University). Purified ADA gp120 was produced in the laboratory of R. Doms, University of Pennsylvania. Fab X5 was expressed in *Escherichia coli* and purified using an anti-human Fab specific affinity column. Deglycosylated gp120 JRFL was expressed in HEK 293S GnTI$^{-/-}$ cells and treated with Endo H (Roche).

**Donors.** The donors identified for this study were selected from the IAVI sponsored study, Protocol G[14]. Eligibility for enrolment into Protocol G was defined as: male or female at least 18 years of age with documented HIV infection for at least three years, clinically asymptomatic at the time of enrolment and not currently receiving antiretroviral therapy. Selection of individuals for monoclonal antibody generation was based on a rank-order high-throughput screening and analytical algorithm[14]. Volunteers were identified as elite neutralizers based on broad and potent neutralizing activity against a cross-clade pseudovirus panel[14].

**Isolation of monoclonal antibodies.** The method for isolating human monoclonal antibodies from memory B cells in circulation has previously been described[2]. Surface IgG$^{+}$ B cells seeded at near-clonal density in 384-well microplates were activated in short-term culture. Supernatants were screened for neutralization activity against 2–4 pseudotyped viruses for which neutralization activity was detected at high titres in the donor serum. Heavy- and light-chain variable regions were isolated from B-cell lysates of selected neutralizing hits by reverse transcription from RNA followed by multiplex PCR amplification using family-specific V-gene primer sets. Amplicons from each lysate were uniquely tagged with multiplex identifier (MID) sequences and 454 sequencing regions (Roche). A normalized pooling of gamma, kappa and lambda chains was performed based on agarose gel image quantitation and the pool was analysed by 454 Titanium sequencing. Consensus sequences of the VH and VL chains were generated using the Amplicon Variant Analyser (Roche) and assigned to specific B-cell culture wells by decoding the MID tags. Selected VH and VL chains were synthesized and cloned in expression vectors with the appropriate IgG1, Ig kappa or Ig lambda constant domain. Monoclonal antibodies were reconstituted by transient transfection in HEK293 cells followed by purification from serum-free culture supernatants.

**PGT antibody expression and purification.** Antibody genes were cloned into an expression vector and transiently expressed with the FreeStyle 293 Expression System (Invitrogen). Antibodies were purified using affinity chromatography (Protein A Sepharose Fast Flow, GE Healthcare) and purity and integrity checked with SDS–PAGE.

**Neutralization assays.** Neutralization by monoclonal antibodies and donor sera was performed by Monogram Biosciences using a single round of replication pseudovirus assay as previously described[31]. Briefly, pseudoviruses capable of a single round of infection were produced by co-transfection of HEK293 cells with a subgenomic plasmid, pHIV-1lucu3, that incorporates a firefly luciferase indicator gene and a second plasmid, pCXAS, which expressed HIV-1 Env libraries or clones. Following transfection, pseudoviruses were harvested and used to infect U87 cell lines expressing co-receptors CCR5 or CXCR4. Pseudovirus neutralization assays using HIV-1JR-CSF alanine mutants are fully described elsewhere[2]. Neutralization activity of monoclonal antibodies against HIV-1JR-CSF alanine mutants was measured using a TZM-BL assay, as described[2]. Kifunensine-treated pseudoviruses were produced by treating 293T cells with 25 μM kifunensine on the day of transfection. Memory B-cell supernatants were screened in a micro-neutralization assay against a cross-clade panel of HIV-1 isolates and $SIV_{MAC239}$ (negative control). This assay was based on the 96-well pseudotyped HIV-1 neutralization assay (Monogram Biosciences) and was modified for screening 15 μl of B-cell culture supernatants in a 384-well format.

**Cell surface binding assays.** Titrating amounts of antibodies were added to HIV-1 Env-transfected 293T cells, incubated for 1 h at 37 °C, washed with FACS buffer, and stained with goat anti-human IgG F(ab′)$_2$ conjugated to phycoerythin (Jackson ImmunoResearch). Binding was analysed using flow cytometry, and binding curves were generated by plotting the mean fluorescence intensity of antigen binding as a function of antibody concentration. For competition assays, titrating amounts of competitor antibodies were added to the cells 30 min before adding biotinylated PGT monoclonal antibodies at a concentration required to give half-maximum effective concentration ($EC_{50}$).

**ELISA assays.** For antigen-binding ELISAs, serial dilutions of monoclonal antibodies were added to antigen-coated wells and binding was probed with alkaline-phosphatase-conjugated goat anti-human IgG F(ab′)$_2$ antibody (Pierce). For competition ELISAs, titrating amounts of competitor monoclonal antibodies were added to gp120-coated ELISA wells and incubated for 30 min before adding biotinylated PGT monoclonal antibodies at a concentration required to give $IC_{70}$. Biotinylated PGT monoclonal antibodies were detected using alkaline-phosphatase-conjugated streptavidin (Pierce) and visualized using p-nitrophenol phosphate substrate (Sigma).

**Glycan microarray analysis.** Monoclonal antibodies were screened on a printed glycan microarray version 5.0 from the CFG as described previously[29]. Antibodies were used at a concentration of 30 μg ml$^{-1}$ and were precomplexed with 15 μg ml$^{-1}$ secondary antibody (goat anti-human-Fc-rPE, Jackson Immunoresearch) before addition to the slide. Complete glycan array data sets for all antibodies can be found at http://www.functionalglycomics.org in the CFG data archive under "cfg_rRequest_2250".

**Oligomannose dendron synthesis.** The oligomannose dendrons ($Man_4D$ and $Man_9D$) were synthesized by Cu(I) catalysed alkyne-azide cycloaddition between azido oligomannose and the second generation of $AB_3$ type alkynyl dendron. Detailed procedures and characterization were previously reported[30].

**Fabrication of gp120 microarray.** NHS-activated glass slides (Nexterion slide H, Schott North American) were printed with robotic pin (Arrayit 946) to deposit gp120 JRFL at concentrations of 750 or 250 μg ml$^{-1}$ in printing buffer (120 mM phosphate, pH 8.5; containing 5% glycerol and 0.01% Tween 20). Twelve replicates were used for each concentration. The printed slides were incubated in relative humidity 75% chamber overnight and treated with blocking solution (superblock blocking buffer in PBS, Thermo) at 25 °C for 1 h. The slides were then rinsed with PBS-T (0.05% Tween 20) and PBS buffer, and centrifuged at 200*g* to remove residual solution from slide surface.

**Oligomannose dendron-gp120 competition assay with monoclonal antibodies.** Serial diluted oligomannose dendrons were mixed with monoclonal antibody (40 μg ml$^{-1}$) in PBS-BT buffer (1% BSA and 0.05% Tween 20 in PBS). The mixtures were applied directly to each sub-array on slide. After incubation in a humidified chamber for 1 h at 25 °C, the slides were rinsed sequentially with PBS-T (0.05% Tween 20 in PBS) and PBS buffer, and then centrifuged at 200*g*. Each sub-array was then stained with Cy3-labelled goat anti-human Fc IgG (7.5 μg ml$^{-1}$ in PBS-BT) for 1 h in a humidified chamber. The slides were then rinsed sequentially with PBS-T and deionized water and centrifuged at 200*g*. The fluorescence of the final arrays was imaged at 10 μm resolution (excitation: 540 nm; emission: 595 nm) with an ArrayWorx microarray reader (Applied Precision).

**Sequence analysis.** Germline genes were predicted using the immunoglobulin sequence alignment tools IMGT/V-QUEST[32] and SoDA2[33]. Clonally related sequences were identified by common germline V-genes and long stretches of identical N-nucleotides.

**Statistics.** Statistical analyses were done with Prism 5.0 for Mac (GraphPad). Viruses that are not neutralized at an $IC_{50}$ or $IC_{90} < 50$ μg ml$^{-1}$ were given a value of 50 μg ml$^{-1}$ for median calculations. For combinations of antibodies, a virus was counted as covered if at least one of the monoclonal antibodies was neutralized depending on individual concentrations ($IC_{50}$). This approach does not take additivity into account and therefore underestimates the neutralization potency of antibody combinations.


31. Richman, D. D., Wrin, T., Little, S. J. & Petropoulos, C. J. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proc. Natl Acad. Sci. USA* **100,** 4144–4149 (2003).
32. Brochet, X., Lefranc, M. P. & Giudicelli, V. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res.* **36,** W503–W508 (2008).
33. Munshaw, S. & Kepler, T. B. SoDA2: a Hidden Markov Model approach for identification of immunoglobulin rearrangements. *Bioinformatics* **26,** 867–872 (2010).


©2011 Macmillan Publishers Limited. All rights reserved

LETTER

doi:10.1038/nature10378

# Parallel evolution of domesticated *Caenorhabditis* species targets pheromone receptor genes

Patrick T. McGrath[1], Yifan Xu[1], Michael Ailion[2], Jennifer L. Garrison[1], Rebecca A. Butcher[3] & Cornelia I. Bargmann[1]

**Evolution can follow predictable genetic trajectories[1], indicating that discrete environmental shifts can select for reproducible genetic changes[2–4]. Conspecific individuals are an important feature of an animal's environment, and a potential source of selective pressures. Here we show that adaptation of two *Caenorhabditis* species to growth at high density, a feature common to domestic environments, occurs by reproducible genetic changes to pheromone receptor genes. Chemical communication through pheromones that accumulate during high-density growth causes young nematode larvae to enter the long-lived but non-reproductive dauer stage. Two strains of *Caenorhabditis elegans* grown at high density have independently acquired multigenic resistance to pheromone-induced dauer formation. In each strain, resistance to the pheromone ascaroside C3 results from a deletion that disrupts the adjacent chemoreceptor genes *serpentine receptor class g* (*srg*)*-36* and *-37*. Through misexpression experiments, we show that these genes encode redundant G-protein-coupled receptors for ascaroside C3. Multigenic resistance to dauer formation has also arisen in high-density cultures of a different nematode species, *Caenorhabditis briggsae*, resulting in part from deletion of an *srg* gene paralogous to *srg-36* and *srg-37*. These results demonstrate rapid remodelling of the chemoreceptor repertoire as an adaptation to specific environments, and indicate that parallel changes to a common genetic substrate can affect life-history traits across species.**

*Caenorhabditis elegans* and many other nematode species evaluate environmental conditions to choose between two alternative developmental trajectories, one leading to rapid reproduction and one leading to arrest in the long-lived, stress-resistant dauer larva stage. High population density, limiting food and high temperature promote dauer larva formation[5] (Fig. 1a), a stage that corresponds to the infectious juvenile stage of parasitic nematodes. Dauer larvae do not feed or reproduce, but can survive under conditions that kill other stages, and respond to environmental improvements by exiting the dauer stage and resuming reproductive development. Although the pheromone cues that signal nematode density are normally integrated with food availability, pheromone accumulation in high-density liquid cultures causes animals to form dauer larvae despite the presence of ample food[6]. Non-reproducing dauer animals would seem to be at a disadvantage relative to those that continue to grow in these conditions. To examine adaptation to high-density culture conditions, we measured dauer formation in two laboratory strains of *C. elegans*, LSJ2 and CC1, that were grown in liquid axenic media for approximately 50 years and 4 years, respectively, before permanent cultures were frozen down[7,8] (Fig. 1b and Methods). Unlike wild-caught strains[9] and the standard laboratory strain N2 (ref. 10), which readily form dauers in response to partially purified N2 dauer pheromone, CC1 and LSJ2 strains formed almost no dauer larvae (Fig. 1c).

N2, LSJ2 and CC1 arose from a common, inbred *C. elegans* ancestor after isolation from the wild (Fig. 1b), so the pheromone resistance of LSJ2 and CC1 strains must result from new mutations that occurred in the laboratory. The genetic basis of dauer pheromone resistance was characterized by generating 94 recombinant inbred lines (RILs) between LSJ2 and N2 (Supplementary Fig. 1) that were genotyped at 176 informative single nucleotide polymorphisms (SNPs) (Supplementary Table 1) identified by whole-genome sequencing of LSJ2 and N2 strains (Supplementary Tables 2 and 3). Initial genetic mapping of dauer formation using N2-derived dauer pheromone preparations and the N2–LSJ2 RILs indicated that the trait was multigenic (data not shown). The active components of dauer pheromone are ascarosides, a group of small molecules with a common sugar scaffold and variable side chains[11–13]. Four individual ascarosides that effectively induced N2 dauer formation (C3, C5, C6 and C9) did not induce dauer formation in LSJ2 or CC1 (Fig. 1c). To simplify trait-mapping, we examined dauer formation in response to individual ascarosides, focusing on the C3 ascaroside, whose receptors and cellular sites of action are unknown. Among 16 RILs exposed to 1 μM C3, eight formed dauers at a rate comparable to N2 and eight formed dauers at a rate comparable to LSJ2 (Fig. 1d). This bimodal distribution indicates the existence of a single locus that confers C3 resistance.

Quantitative trait locus (QTL) mapping using these sixteen RILs identified a single region on the X chromosome that correlated with the C3 response (Fig. 1e). Mapping of the X-linked C3 resistance locus was verified by creating a near-isogenic line (NIL) with the candidate region from LSJ2 introgressed into an N2 background by ten generations of backcrossing (Fig. 1f). The LSJ2–N2 NIL was resistant to dauer formation induced by C3 ascaroside across a broad range of concentrations (Fig. 1g). Unlike the parental LSJ2 strain, the LSJ2–N2 NIL formed dauer larvae in the presence of three other ascarosides (Fig. 1g). These results identify an X-linked C3-resistance locus as one of several loci that confer pheromone resistance on LSJ2.

To identify the genetic changes in LSJ2 associated with C3 resistance, we sequenced the LSJ2 and N2 strains and identified all fixed polymorphisms between the two strains (Supplementary Tables 2 and 3). The region of the X chromosome associated with C3 resistance included four SNPs in intronic or intergenic regions and a deletion of 4,906 base pairs (bp) in the LSJ2 strain that disrupts two predicted G-protein-coupled receptor genes (*srg-36* and *srg-37*) (Fig. 2a). A genomic clone from the N2 strain that contains both *srg-36* and *srg-37* fully rescued C3 resistance when introduced into the LSJ2–N2 NIL strain, indicating that this deletion causes the C3 resistance associated with the X-linked QTL (Fig. 2b). Notably, *srg-36* and *srg-37* were also disrupted by a 6,795-bp deletion in the CC1 strain (Fig. 2a). The deletions in CC1 and LSJ2 have different breakpoints, indicating that they occurred independently. To ask whether deletion of *srg-36* and *srg-37* also caused resistance to C3 in CC1, the region surrounding the *srg-36* and *srg-37* deletion was introgressed from CC1 into N2 to make a CC1–N2 NIL strain (Fig. 1f). The CC1–N2 NIL was resistant to dauer formation induced by C3 ascaroside (Fig. 2b) and its C3-resistance phenotype was rescued by a transgene covering the *srg-36* and *srg-37* genomic regions (Fig. 2b). The CC1–N2 NIL readily formed dauers in response to other ascarosides (Supplementary Fig. 2), indicating that additional genetic mutations contribute to pheromone resistance in

[1]Howard Hughes Medical Institute, Laboratory of Neural Circuits and Behavior, The Rockefeller University, New York, New York 10065, USA. [2]Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA. [3]Department of Chemistry, University of Florida, Gainesville, Florida 32611, USA.



©2011 Macmillan Publishers Limited. All rights reserved

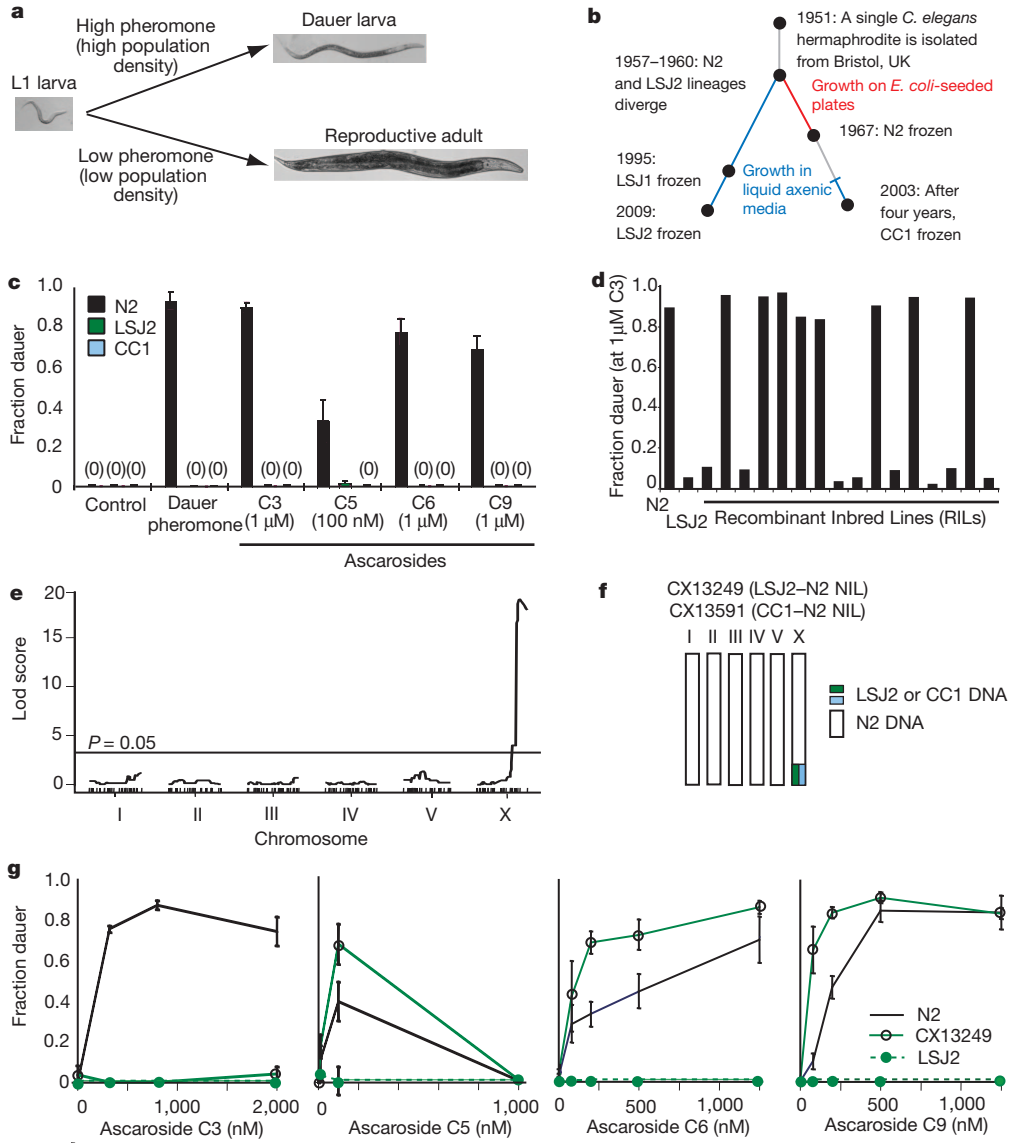



**Figure 1 | Strains of *C. elegans* cultivated in liquid are resistant to dauer pheromones.** **a**, The developmental decision between reproductive growth and dauer larva formation is regulated by temperature, food and population density. Population density is assessed by the release and sensation of ascarosides including C3, C5, C6 and C9. **b**, History of the *C. elegans* strains N2, LSJ1, LSJ2 and CC1 (see Methods). **c**, Dauer formation of N2, LSJ2 and CC1 in response to crude dauer pheromone or synthetic ascarosides. **d**, Dauer formation in response to synthetic C3 ascaroside. **e**, QTL mapping of C3 resistance. **f**, Schematic of NILs with a small region from LSJ2 or CC1 introgressed into N2. **g**, Dauer formation in N2, LSJ2 and CX13249 strains. Error bars represent s.e.m.

CC1. The existence of independent C3-resistance mutations affecting *srg-36* and *srg-37* in LSJ2 and CC1 provides strong genetic evidence linking these two chemoreceptors to dauer formation.

To determine which of the two predicted genes is associated with C3 sensitivity, we introduced *srg-36* and *srg-37* complementary DNAs (cDNAs) with their respective upstream regions into the C3-resistant LSJ2–N2 NIL strain (Fig. 2a). Transgenic strains expressing either of the two cDNAs formed dauer larvae in response to C3, although the *srg-37* transgene was less active than the *srg-36* transgene (Fig. 2b). An *srg-37* genomic fragment also rescued dauer formation (Fig. 2b). These results indicate that the *srg-36* and *srg-37* genes are at least partially redundant; either can support dauer formation in response to C3 ascaroside.

The expression patterns of *srg-36* or *srg-37* were inferred from bicistronic transcripts expressing green fluorescent protein (GFP) downstream of the *srg-36* or *srg-37* promoter and cDNA. These *srg-36* and *srg-37* reporter transgenes rescued C3-induced dauer formation (Fig. 2b), and were most strongly and consistently expressed in the ASI chemosensory neurons, with weak or inconsistent expression in a few other neurons (Fig. 2c). Reporters for *srg-36* and *srg-37* were robustly expressed during the L1 stage when the dauer decision is made (Fig. 2c). The ASI neurons are primary regulators of dauer formation[14], and are therefore plausible sites of *srg-36* and *srg-37* action. An *srg-36* cDNA driven by the ASI-selective *srg-47* promoter rescued C3-induced dauer formation in the LSJ2–N2 NIL, but expression of *srg-36* in AFD or ASE sensory neurons did not (Fig. 2b and Supplementary Fig. 3). These results are consistent with the hypothesis that *srg-36* acts in ASI to sense ascaroside C3 (Supplementary Fig. 4).

The subcellular localization of SRG-36 was examined by fusing GFP to the *srg-36* cDNA and expressing the hybrid gene from an ASI-specific promoter. This fusion protein was primarily localized in the sensory cilia of ASI (Fig. 3a), indicating a sensory function for SRG-36. The selective association of *srg-36* and *srg-37* with C3 responsiveness, and not with responsiveness to other ascarosides, suggested that they might encode C3 receptors. We tested this hypothesis by a gain-of-function experiment in which the *srg-36* cDNA or the *srg-37* cDNA



©2011 Macmillan Publishers Limited. All rights reserved

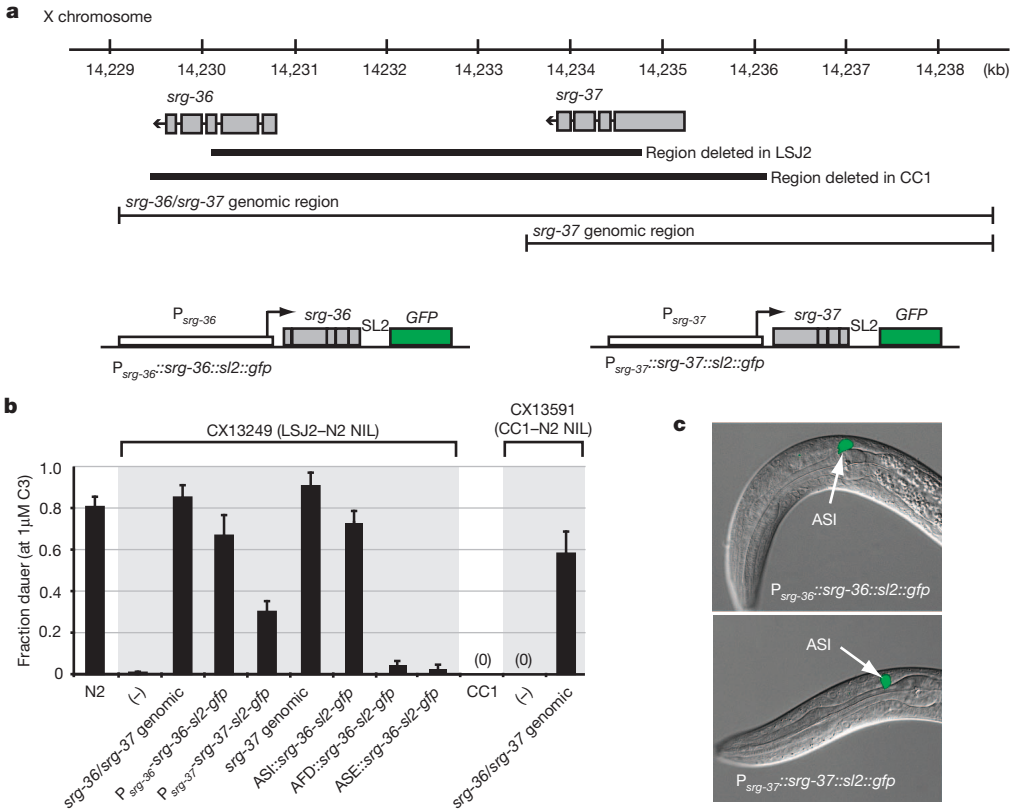



**Figure 2 | Resistance to C3 ascaroside is caused by deletion of two *srg* genes.** **a**, Genomic region surrounding *srg-36* and *srg-37* on the X chromosome, deletion breakpoints in LSJ2 and CC1 strains, fragments used for transgenic rescue, and design of bicistronic fusion genes. **b**, Transgenic rescue of dauer formation in response to C3 ascaroside. NIL strains used as recipients for rescue are shown in Fig. 1f. The ASI promoter was *srg-47* (Supplementary Fig. 3), the AFD promoter was *gcy-8* and the ASE promoter was *flp-6*. Error bars represent s.e.m. **c**, Expression of GFP from bicistronic fusion genes for *srg-36* and *srg-37* in L1 larvae, showing predominant expression in ASI sensory neurons.

was expressed in ASH neurons, a pair of polymodal nociceptive neurons that direct rapid avoidance behaviour[15]. Unlike control animals, animals expressing the ASH::*srg-36* or ASH::*srg-37* transgene reversed rapidly in response to 1 μM C3 in an acute-avoidance assay (Fig. 3b). Neither ASH::*srg-36*, ASH::*srg-37*, nor control animals responded strongly to 1 μM C6 (Fig. 3b). These results demonstrate that expression of SRG-36 or SRG-37 in ASH is sufficient for C3-specific behavioural responses.

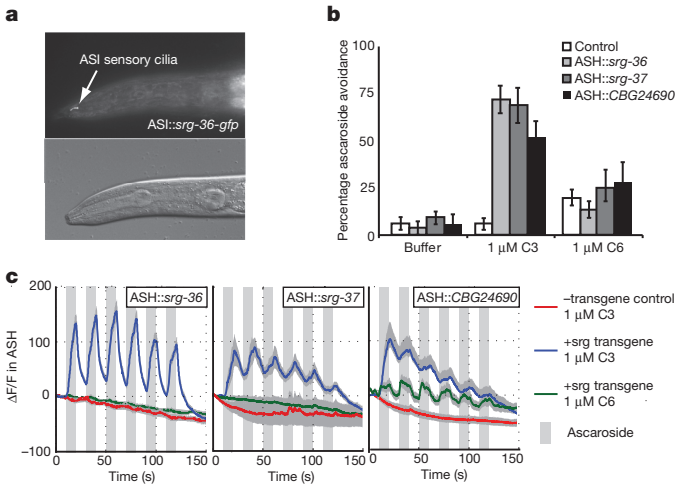



**Figure 3 | The *srg* genes encode ascaroside receptors.** **a**, Localization of SRG-36::GFP to ASI cilia (L4 animal). **b**, Ascaroside avoidance behaviours of animals with ectopic expression of *srg-36*, *srg-37* or *CBG24690* (shown in Fig. 4) in the ASH nociceptive neurons. Error bars represent s.e.m. **c**, Ascaroside-induced $Ca^{2+}$ transients in ASH neurons that ectopically express *C. elegans srg-36* or *srg-37*, or *C. briggsae CBG24690*, in ASH. Grey bars indicate the presence of C3 or C6 ascaroside, shading indicates s.e.m., $n \geq 10$ animals per condition. $Ca^{2+}$ was monitored using the genetically-encoded calcium sensor GCaMP3.0 (ref. 30). ΔF/F, percentage fluorescence change (baseline fluorescence = 100%).

The ASH neurons respond to repulsive stimuli with increases in intracellular calcium that can be monitored using genetically-encoded calcium indicators[16]. Animals expressing *srg-36* or *srg-37* in ASH showed rapid, reliable $Ca^{2+}$ increases in response to 1 μM C3 ascaroside, but not to 1 μM C6 ascaroside (Fig. 3c); control animals did not respond to either C3 or C6. These results indicate that SRG-36 and SRG-37 are chemoreceptors (or subunits of chemoreceptors) that sense the C3 ascaroside. Although *srg-36* and *srg-37* are normally expressed in ASI, ASI neurons did not respond to C3 with calcium transients (data not shown). Little is known about pheromone signalling pathways in ASI, so the reason for this negative result is unclear.

LSJ2 was originally propagated in the Dougherty laboratory in the 1950s and 1960s to study nutrient requirements for nematode growth. A strain of *C. briggsae*, DR1690, that was grown in the Dougherty laboratory under the same conditions as LSJ2 also acquired resistance to dauer pheromone[17] (Fig. 4a). *C. briggsae* and *C. elegans* are estimated to have diverged 20–30 million years ago[18]. The *C. briggsae* genome encodes several genes closely related to *srg-36* and *srg-37*, but does not have one-to-one orthologues of these genes (Fig. 4b). Comparing genomic DNA sequences from DR1690 with the reference *C. briggsae* strain AF16, we discovered a 33-kilobase deletion in DR1690 that disrupts one of the *srg* paralogs, *CBG24690*, and six other genes (Fig. 4c). To determine whether this deletion affects pheromone responses, we created a NIL with the *CBG24690* deletion introgressed into the AF16 reference background (DR1690–AF16 NIL). As previously



©2011 Macmillan Publishers Limited. All rights reserved

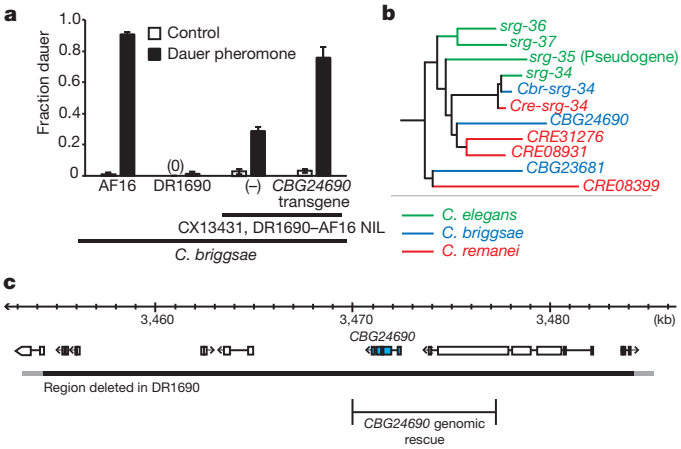



**Figure 4 | Evolutionary conservation of *srg* function.** **a**, Rescue of *C. briggsae* dauer formation in response to partially purified dauer pheromone, mediated by genomic fragments containing the *CBG24690* gene. CX13431 is a near-isogenic line containing the *CBG24690* deletion from DR1690, introgressed into the AF16 background. Error bars represent s.e.m. **b**, Schematic of genes closely related to *srg-36* and *srg-37* from *C. elegans*, *C. briggsae* and *C. remanei* (adapted from ref. 19). **c**, *CBG24690* genomic region from the AF16 *C. briggsae* reference strain, and location of a large deletion in the DR1690 *C. briggsae* strain that was cultivated for an extended period in liquid axenic media.

reported[17], AF16 readily formed dauers in response to *C. elegans* dauer pheromone, whereas DR1690 animals were resistant to dauer pheromone (Fig. 4a). The *C. briggsae* DR1690–AF16 NIL formed dauers at an intermediate level compared to these two strains (Fig. 4a), indicating that this region contains one of several mutations that contribute to pheromone resistance in DR1690. The DR1690–AF16 NIL was also resistant to purified ascaroside C3 compared to the parental AF16 strain (data not shown). A transgene covering the *CBG24690* genomic region rescued dauer formation in the pheromone-resistant DR1690–AF16 NIL strain (Fig. 4a). These results demonstrate that the *CBG24690 srg* gene contributes to pheromone-induced dauer formation in *C. briggsae*.

To investigate whether *CBG24690* also encodes an ascaroside receptor, we expressed a *CBG24690* cDNA in the *C. elegans* ASH neurons. Animals expressing the ASH::*CBG24690* transgene reversed rapidly when presented with 1 μM C3, but not when presented with 1 μM C6 (Fig. 3b). Animals expressing the *CBG24690 srg* gene in ASH also showed rapid $Ca^{2+}$ increases in response to 1 μM C3 (Fig. 3c). Unlike animals expressing *srg-36* or *srg-37* in ASH, animals expressing *CBG24690* in ASH showed weaker but reliable $Ca^{2+}$ increases in response to the related ascaroside C6 (Fig. 3c).

Our results indicate that *srg-36* and *srg-37*, two members of a large nematode-specific family of G-protein-coupled receptors[19], encode redundant receptors for the ascaroside C3. The *srg* gene family is distinct from the *srbc* gene family that was previously implicated in sensing ascarosides C6 and C9 (ref. 20), indicating that at least two of the seven chemoreceptor superfamilies of *C. elegans* can detect ascaroside pheromones. Chemoreceptors are among the fastest-evolving genes in metazoan genomes. They come from entirely different protein families in vertebrates, insects and nematodes, and change rapidly between species[21]: only half of the chemoreceptors in *C. elegans* and *C. briggsae* are one-to-one orthologue pairs[19]. Despite the rapid evolution of these genes, the function of *srg*-like genes in pheromone detection has been conserved since *C. briggsae* and *C. elegans* diverged: *srg-36* and *srg-37* in *C. elegans*, and *CBG24690* in *C. briggsae*, each sense C3 ascaroside to induce dauer formation. Some differences between species exist, however, because *CBG24690* also senses C6 ascaroside at concentrations that are not sensed by *srg-36* and *srg-37*. Differences in pheromone production and sensation by different *Caenorhabditis* species may allow both species-specific discrimination and general detection of *Caenorhabditis* species in the vicinity, as is observed with quorum-sensing systems in bacteria[22].

These results demonstrate a reproducible change in the chemoreceptor repertoire in response to a discrete environmental shift. During high-density growth in the laboratory, resistance to ascaroside pheromones arose independently in *C. elegans* LSJ2 and CC1, and in *C. briggsae* DR1690, through changes in related *srg* genes. Although numerous single-gene mutations can convey resistance to dauer formation in *C. elegans*[23], deletion of *srg* genes seems to be a favoured route to pheromone resistance in mixed populations. It is possible that specific features of the chromosomal region surrounding *srg-36* and *srg-37* predispose this region to deletion mutations, but these features would also need to be found near the *CBG24690* gene in *C. briggsae*. Alternatively, the spectrum of potential dauer-defective mutants may be constrained because the known single-gene mutants that are resistant to dauer formation have pleiotropic effects on sensory biology, stress resistance and starvation responses that would reduce their fitness in mixed cultures[23]. Global analysis of functional genetic variants indicates that evolutionarily relevant mutations are not randomly distributed, but rather cluster in specific genetic loci, or hotspot genes[1]. These observations indicate that genetic trajectories during evolution are constrained and that adaptation can, at least to some extent, be predictable. One class of known adaptive genes are input–output genes, developmental regulators with complex *cis*-regulatory motifs that provide a molecular substrate that allows sculpting of developmental patterns[2,24]. The genes *srg-36* and *srg-37* seem to fall into a second class of adaptive genes, including opsin genes and taste receptors[25,26]: sensory receptors whose diversity allows circumscribed adaptation to environmental changes without pleiotropic effects.

## METHODS SUMMARY

The LSJ2 strain and the N2 laboratory strain are descended from one ancestral hermaphrodite isolated by W. Nicholas. LSJ2 was grown continuously in liquid axenic media starting in about 1957 at the Kaiser Foundation Research Institute, The University of California, Berkeley, and San Jose State University, until a sample was frozen in 2009.

Dauer formation assays were performed with crude or synthesized ascarosides as described[11]. Values report the average fraction of dauer animals 72 h after eggs were laid on assay plates. With the exception of the mapping experiments, each strain and condition was tested in a minimum of five independent assays.

RILs were generated from reciprocal crosses between LSJ2 and an N2-derived strain, and were inbred for ten generations. Two laboratory-derived polymorphisms that modify the *npr-1* and *glb-5* genes in N2 affect many *C. elegans* behaviours[7,27]; to eliminate their effects, the cross was initiated with a strain containing 99% N2 DNA but the ancestral alleles of *npr-1* and *glb-5* from the CB4856 strain (Supplementary Fig. 1). These RILs were genotyped at 192 SNPs between LSJ2 and N2. The fraction of animals forming dauers in response to 1 μM C3 ascaroside was used as a phenotype for nonparametric QTL mapping.

The genes *srg-36*, *srg-37* and *CBG24690* were ectopically expressed in ASH using the *sra-6* promoter. Vehicle control, 1 μM of C3 or 1 μM of C6 were dissolved in M13 buffer and presented to animals using the drop test[28]. Each animal was scored three times for the ability to reverse in response to the stimulus. At least 50 animals were scored blindly for each strain and condition.

ASH imaging was performed in a custom-designed microfluidic device[29]. The genetically-encoded calcium indicator GCaMP3.0 (ref. 30) was expressed in ASH using the *sra-6* promoter.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 28 February; accepted 20 July 2011.**
**Published online 17 August 2011.**

1. Stern, D. L. & Orgogozo, V. Is genetic evolution predictable? *Science* **323,** 746–751 (2009).
2. Chan, Y. F. *et al.* Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* **327,** 302–305 (2010).
3. Protas, M. E. *et al.* Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genet.* **38,** 107–111 (2006).
4. Woods, R., Schneider, D., Winkworth, C. L., Riley, M. A. & Lenski, R. E. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **103,** 9107–9112 (2006).



©2011 Macmillan Publishers Limited. All rights reserved

5. Golden, J. W. & Riddle, D. L. The *Caenorhabditis elegans* dauer larva: developmental effects of pheromone, food, and temperature. *Dev. Biol.* **102,** 368–378 (1984).
6. Stiernagle, T. Maintenance of *C. elegans*. *WormBook* (ed. The *C. elegans* Research Community) doi:10.1895/wormbook. 1.101.1 (11 February 2006); available at http://www.wormbook.org.
7. McGrath, P. T. *et al.* Quantitative mapping of a digenic behavioral trait implicates globin variation in *C.elegans* sensory behaviors. *Neuron* **61,** 692–699 (2009).
8. Szewczyk, N. J., Kozak, E. & Conley, C. A. Chemically defined medium and *Caenorhabditis elegans*. *BMC Biotechnol.* **3,** 19 (2003).
9. Viney, M. E., Gardner, M. P. & Jackson, J. A. Variation in *Caenorhabditis elegans* dauer larva formation. *Dev. Growth Differ.* **45,** 389–396 (2003).
10. Golden, J. W. & Riddle, D. L. A pheromone influences larval development in the nematode *Caenorhabditis elegans*. *Science* **218,** 578–580 (1982).
11. Butcher, R. A., Ragains, J. R., Kim, E. & Clardy, J. A potent dauer pheromone component in *Caenorhabditis elegans* that acts synergistically with other components. *Proc. Natl Acad. Sci. USA* **105,** 14288–14292 (2008).
12. Jeong, P. Y. *et al.* Chemical structure and biological activity of the *Caenorhabditis elegans* dauer-inducing pheromone. *Nature* **433,** 541–545 (2005).
13. Srinivasan, J. *et al.* A blend of small molecules regulates both mating and development in *Caenorhabditis elegans*. *Nature* **454,** 1115–1118 (2008).
14. Bargmann, C. I. & Horvitz, H. R. Control of larval development by chemosensory neurons in *Caenorhabditis elegans*. *Science* **251,** 1243–1246 (1991).
15. Kaplan, J. M. & Horvitz, H. R. A dual mechanosensory and chemosensory neuron in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **90,** 2227–2231 (1993).
16. Hilliard, M. A. *et al. In vivo* imaging of *C.elegans* ASH neurons: cellular response and adaptation to chemical repellents. *EMBO J.* **24,** 63–72 (2005).
17. Fodor, A., Riddle, D. L., Nelson, F. K. & Golden, J. W. Comparison of a new wild-type *Caenorhabditis briggsae* with laboratory strains of *C.briggsae* and *C. elegans*. *Nematologica* **29,** 203–216 (1983).
18. Cutter, A. D. Divergence times in *Caenorhabditis* and *Drosophila* inferred from direct estimates of the neutral mutation rate. *Mol. Biol. Evol.* **25,** 778–786 (2008).
19. Thomas, J. H. & Robertson, H. M. The *Caenorhabditis* chemoreceptor gene families. *BMC Biol.* **6,** 42 (2008).
20. Kim, K. *et al.* Two chemoreceptors mediate developmental effects of dauer pheromone in *C.elegans*. *Science* **326,** 994–998 (2009).
21. Nei, M., Niimura, Y. & Nozawa, M. The evolution of animal chemosensory receptor gene repertoires: roles of chance and necessity. *Nature Rev. Genet.* **9,** 951–963 (2008).
22. Waters, C. M. & Bassler, B. L. Quorum sensing: cell-to-cell communication in bacteria. *Annu. Rev. Cell Dev. Biol.* **21,** 319–346 (2005).
23. Hu, P. J. Dauer. *WormBook* (ed. The *C. elegans* Research Community) doi:10.1895/wormbook.1.144.1 (8 August 2007); available at http://www.wormbook.org.
24. Sucena, E., Delon, I., Jones, I., Payre, F. & Stern, D. L. Regulatory evolution of *shavenbaby/ovo* underlies multiple cases of morphological parallelism. *Nature* **424,** 935–938 (2003).
25. Kim, U. K. *et al.* Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. *Science* **299,** 1221–1225 (2003).
26. Yokoyama, S. Molecular evolution of vertebrate visual pigments. *Prog. Retin. Eye Res.* **19,** 385–419 (2000).
27. de Bono, M. & Bargmann, C. I. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C.elegans*. *Cell* **94,** 679–689 (1998).
28. Hilliard, M. A., Bargmann, C. I. & Bazzicalupo, P. C. *elegans* responds to chemical repellents by integrating sensory inputs from the head and the tail. *Curr. Biol.* **12,** 730–734 (2002).
29. Chronis, N., Zimmer, M. & Bargmann, C. I. Microfluidics for *in vivo* imaging of neuronal and behavioral activity in *Caenorhabditis elegans*. *Nature Methods* **4,** 727–731 (2007).
30. Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* **6,** 875–881 (2009).


**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank N. Lu for the LSJ2 strain, M. Rockman for a *qglR1(X,CB4856>N2)* introgression strain, J. Ragains for ascaroside synthesis, H. Hang for assistance in purifying pheromones and S. Dewell, K. Foster, N. Ringstad, A. Bendesky, Y. Saheki, M. Zimmer, S. Crosson, E. Feinberg, E. Toro, M. Rockman and L. Kruglyak for comments and advice. P.T.M. was funded by a Damon Runyon Fellowship. Y.X. was supported by Medical Scientist Training Program (MSTP) grant GM07739 and a Paul and Daisy Soros Fellowship. J.L.G. was an HHMI fellow of the Helen Hay Whitney Foundation and is funded by National Institutes of Health (NIH) K99 GM092859. R.A.B. is supported by R00GM87533. C.I.B. is an investigator of the Howard Hughes Medical Institute. This work was supported by the HHMI.

**Author Contributions** P.T.M. and C.I.B. designed and interpreted experiments and wrote the paper. P.T.M. performed all genetic, molecular and behavioural experiments, Y.X. conducted calcium imaging experiments, M.A. identified the dauer-formation defect in the LSJ2 lineage, R.A.B. characterized and synthesized ascarosides and J.L.G. contributed reagents.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.I.B. (cori@rockefeller.edu).



©2011 **Macmillan Publishers Limited. All rights reserved**

## METHODS

Strains were cultivated at 22 °C on agar plates seeded with *E. coli* strain OP50.

LSJ2 is a sister strain to the standard N2 laboratory strain. Both LSJ2 and N2 are descended from a single animal isolated by W. Nicholas from a mushroom compost culture provided by L. Staniland in Bristol, England. The strain was transferred to E. Dougherty's laboratory at the Kaiser Foundation Research Institute in the 1950s, and in the late 1950s it separated into two substrains. One of these substrains was mistakenly believed to be *C. briggsae* and represents the LSJ2 lineage. S. Brenner received a cultivar of the second substrain from E. Dougherty in 1964; this cultivar became N2. The LSJ2 lineage was continuously cultivated in liquid axenic media at the University of California, Berkeley and at San Jose State University thereafter. In 1995, a cultivar of the strain was sent to the *Caenorhabditis* genetics centre and frozen to become LSJ1. In 2009, N. Lu from San Jose State University provided a second cultivar of the strain that had been grown for an additional 14 years in axenic media; this strain is designated LSJ2.

Other strains used in this study are: N2, CC1, LSJ1, MY14, AF16, DR1690, CX12311 *kyIR1(V, CB4856>N2)*; *qgIR1(X, CB4856>N2)*, CX13249 *kyIR88(X, LSJ2>N2)*, CX13330 *kyIR88(X, LSJ2>N2)*; *kyEx3927 (srg-36/srg-37 genomic region + $P_{elt\text{-}2}$::gfp)*, CX13331 *kyIR88(X, LSJ2>N2); kyEx3928 (srg-36/srg-37 genomic region + $P_{elt\text{-}2}$::gfp)*, CX13332 *kyIR88(X, LSJ2>N2); kyEx3929 (srg-37 genomic region + $P_{elt\text{-}2}$::gfp)*, CX13333 *kyIR88(X, LSJ2>N2); kyEx3930 (srg-37 genomic region + $P_{elt\text{-}2}$::gfp)*, CX13334 *kyIR88(X, LSJ2>N2); kyEx3931 (srg-37 genomic region + $P_{elt\text{-}2}$::gfp)*, CX13335 *kyIR88(X, LSJ2>N2); kyEx3932 ($P_{srg\text{-}36}$::srg-36:sl2:gfp + $P_{elt\text{-}2}$::gfp)*, CX13336 *kyIR88(X, LSJ2>N2); kyEx3933 ($P_{srg\text{-}36}$::srg-36:sl2:gfp + $P_{elt\text{-}2}$::gfp)*, CX13337 *kyIR88(X, LSJ2>N2); kyEx3934 ($P_{srg\text{-}36}$::srg-36:sl2:gfp + $P_{elt\text{-}2}$::gfp)*, CX13338 *kyIR88(X, LSJ2>N2); kyEx3935 ($P_{srg\text{-}37}$::srg-37:sl2:gfp + $P_{elt\text{-}2}$::gfp)*, CX13339 *kyIR88(X, LSJ2>N2); kyEx3936 ($P_{srg\text{-}37}$::srg-37:sl2:gfp + $P_{elt\text{-}2}$::gfp)*, CX13340 *kyIR88(X, LSJ2>N2); kyEx3937 ($P_{srg\text{-}37}$::srg-37:sl2:gfp + $P_{elt\text{-}2}$::gfp)*, CX13431 *kyIR94(X, DR1690> AF16)*, CX13591 *kyIR95(X, CC1>N2)*, CX13592 *kyIR95(X, CC1>N2); kyEx4118 (srg-36/srg-37 genomic region + $P_{elt\text{-}2}$::gfp)*, CX13593 *kyIR95(X, CC1>N2); kyEx4119 (srg-36/srg-37 genomic region + $P_{elt\text{-}2}$::gfp)*, CX13594 *kyIR95(X, CC1>N2); kyEx4120 (srg-36/srg-37 genomic region + $P_{elt\text{-}2}$::gfp)*, CX13685 *kyEx2865 ($P_{sra\text{-}6}$::GCaMP3.0 + $P_{ofm\text{-}1}$::gfp); kyEx4171 ($P_{sra\text{-}6}$::srg-36 + $P_{ofm\text{-}1}$::rfp)*, CX13686 *kyEx2865 ($P_{sra\text{-}6}$::GCaMP3.0 + $P_{ofm\text{-}1}$::gfp); kyEx4172 ($P_{sra\text{-}6}$::srg-36 + $P_{ofm\text{-}1}$::rfp)*, CX13687 *kyEx2865 ($P_{sra\text{-}6}$::GCaMP3.0 + $P_{ofm\text{-}1}$::gfp); kyEx4173 ($P_{sra\text{-}6}$::srg-36 + $P_{ofm\text{-}1}$::rfp)*, CX13603 *kyIR88(X, LSJ2>N2); kyEx4125 ($P_{str\text{-}3}$::srg-36::gfp + $P_{elt\text{-}2}$::gfp)*, CX13739 *kyIR94(X, DR1690> AF16); kyEx4201 (CBG24690 genomic region + $P_{myo\text{-}2}$::mcherry)*, CX13740 *kyIR94(X, DR1690> AF16); kyEx4202 (CBG24690 genomic region + $P_{myo\text{-}2}$::mcherry)*, CX13741 *kyIR94(X, DR1690> AF16); kyEx4203 (CBG24690 genomic region + $P_{myo\text{-}2}$::mcherry)*, CX13977 *kyIR88(X, LSJ2>N2); kyEx4316 ($P_{srg\text{-}47}$::srg-36::sl2::gfp + $P_{ofm\text{-}1}$::rfp)*, CX13978 *kyIR88(X, LSJ2>N2); kyEx4317 ($P_{srg\text{-}47}$::srg-36::sl2::gfp + $P_{ofm\text{-}1}$::rfp)*, CX13979 *kyIR88(X, LSJ2>N2); kyEx4318 ($P_{srg\text{-}47}$::srg-36::sl2::gfp + $P_{ofm\text{-}1}$::rfp)*, CX13980 *kyIR88(X, LSJ2>N2); kyEx4319 ($P_{gcy\text{-}8}$::srg-36::sl2::gfp + $P_{ofm\text{-}1}$::rfp)*, CX13981 *kyIR88(X, LSJ2>N2); kyEx4320 ($P_{gcy\text{-}8}$::srg-36::sl2::gfp + $P_{ofm\text{-}1}$::rfp)*, CX13982 *kyIR88(X, LSJ2>N2); kyEx4321 ($P_{gcy\text{-}8}$::srg-36::sl2::gfp + $P_{ofm\text{-}1}$::rfp)*, CX13983 *kyEx2865 ($P_{sra\text{-}6}$::GCaMP3.0 + $P_{ofm\text{-}1}$::gfp); kyEx4322 ($P_{sra\text{-}6}$::srg-37.c + $P_{ofm\text{-}1}$::rfp)*, CX13984 *kyEx2865 ($P_{sra\text{-}6}$::GCaMP3.0 + $P_{ofm\text{-}1}$::gfp); kyEx4323 ($P_{sra\text{-}6}$::srg-37.c + $P_{ofm\text{-}1}$::rfp)*, CX13985 *kyEx2865 ($P_{sra\text{-}6}$::GCaMP3.0 + $P_{ofm\text{-}1}$::gfp); kyEx4324 ($P_{sra\text{-}6}$::srg-37.c + $P_{ofm\text{-}1}$::rfp)*, CX13986 *kyEx2865 ($P_{sra\text{-}6}$::GCaMP3.0 + $P_{ofm\text{-}1}$::gfp); kyEx4325 ($P_{sra\text{-}6}$::CBG24690 + $P_{ofm\text{-}1}$::rfp)*, CX13987 *kyEx2865 ($P_{sra\text{-}6}$::GCaMP3.0 + $P_{ofm\text{-}1}$::gfp); kyEx4326 ($P_{sra\text{-}6}$::CBG24690 + $P_{ofm\text{-}1}$::rfp)*, CX13988 *kyEx2865 ($P_{sra\text{-}6}$::GCaMP3.0 + $P_{ofm\text{-}1}$::gfp); kyEx4327 ($P_{sra\text{-}6}$::CBG24690 + $P_{ofm\text{-}1}$::rfp)*, CX14023 *kyIR88(X, LSJ2>N2); kyEx4342 ($P_{flp\text{-}6}$::srg-36::sl2::gfp + $P_{ofm\text{-}1}$::rfp)*, CX14024 *kyIR88(X, LSJ2>N2); kyEx4343 ($P_{flp\text{-}6}$::srg-36::sl2::gfp + $P_{ofm\text{-}1}$::rfp)*.

**Dauer formation assays.** Dauer plates contained 1 µl (for *C. elegans*) or 25 µl (for *C. briggsae*) of crude *C. elegans* dauer pheromone, or 80 nM–2 µM ascarosides (synthesized as previously described[11,31,32]), in NGM agar without peptone (2.2% Noble Agar, 5 µg ml$^{-1}$ cholesterol, 15 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgSO_4$ and 25 mM $KPO_4$). For *C. elegans,* 20 µl of heat-killed *E. coli* OP50 bacteria (10 µg ml$^{-1}$) were added to each plate, five adult animals were picked onto the plate, allowed to lay eggs for 4 h and then removed. Plates were incubated at 25 °C for 72 h before being scored for dauers, identified by a thin body morphology and non-pumping pharynx. At least five plates were assayed for each strain and condition. For *C. briggsae,* OP50 lawns killed with 50 mg ml$^{-1}$ streptomycin were used, because otherwise animals crawled off the heat-killed bacterial lawn and died. Higher levels of pheromone were required to induce dauer formation on the streptomycin-killed bacteria.

Crude dauer pheromone was purified from 2 l of N2 cultured in S basal medium with HB101 bacteria for 11 days. Supernatants were clarified by centrifugation, further filtered through a Buchner filter funnel (medium frit, Chemglass) under vacuum, then filtered through 0.2 µm PES membranes (Nalgene), concentrated using a rotary evaporator and lyophilized. Solids were extracted three times with 100% ethanol (100 ml each), and the eluents were combined and concentrated using a rotary evaporator to yield 5 ml of crude dauer pheromone (stored at −20 °C).

**LSJ2 and N2 sequencing and analysis.** Genomic DNA was isolated from seven strains: LSJ2, LSJ1 (a sample from the LSJ2 lineage frozen in 1995), MY14 (a wild strain used as an outgroup) and four EMS-mutagenized N2-derived strains. Genomic DNA (10 µg) was provided to the Rockefeller Genomics Resource Center for sequencing. DNA samples were processed using the gDNA paired-end sample preparation kit from Illumina, and sequencing was performed using a GAII instrument.

**SNP analysis.** Sequencing reads with an average quality score above 27 (Sanger format) were aligned to the WS195 *C. elegans* reference sequence and used to identify SNPs using the MAQ software suite (version 0.7.1 easyrun command, using default settings)[33]. The final filtered SNPs (the cns.final.snp file) for each strain were further analysed using custom software that analysed the number of reference and mutant reads that were present for the polymorphisms in all the sequenced strains. Many of the predicted SNPs, both in LSJ2 and in N2, were supported by reads that matched both the reference N2 nucleotide and a mutant nucleotide. These 'heterozygous' SNPs could represent heterozygous alleles maintained by balancing selection, but different levels of coverage of the two reads indicates that these apparent SNPs are actually alignment errors.

To be considered a true polymorphism between the LSJ2 and the N2 strains, we required at least 90% of the reads from the LSJ2 sequencing to be mutant, and fewer than 10% of the reads from the N2-derived strains to be mutant. A total of 223 SNPs passed these criteria. Using MY14 as an outgroup, the SNPs were then classified into the LSJ2 branch if fewer then 10% of the reads from the MY14 sequencing were mutant, and into the N2 branch if more then 75% of the reads from the MY14 sequencing were mutant. Eight SNPs could not be classified because there were no reads from the MY14 sequencing. We broke down the LSJ2 lineage further into mutations occurring before and after 1995, using sequence from the LSJ1 strain. If more then 90% of the reads from LSJ1 supported the mutant read, then the SNP was classified as occurring before 1995. If fewer then 25% of the reads from LSJ1 supported the mutant read, then the SNP was classified as occurring after 1995. One SNP could not be classified.

A recent whole-genome sequencing report indicated a substantially higher level of mutation between N2 and LSJ1 than we detected here, with 877 SNPs instead of 171 (ref. 34). Fourteen SNPs predicted by that analysis, but not by this one, were examined by PCR and Sanger sequencing of N2 and LSJ1; 13 of the 14 were not confirmed and one was ambiguous. If these SNPs are representative, ~80% of the SNPs in the previous report are either miscalled bases or SNPs specific to that laboratory's strains.

**Indel analysis.** We created a custom algorithm to identify insertions and deletions (indels) in LSJ2 with respect to the N2 reference. Because MAQ does not use gapped alignment for aligning single-end reads to the reference sequence, we reasoned that most reads covering an insertion or deletion would be unaligned by the MAQ software. We identified regions of low coverage (defined as <12 reads) using custom software and identified any reads unaligned by MAQ with partial matches (defined as reads with 18 contiguous matches) in these regions. We then realigned the partial reads, considering all possible 1-bp insertions and deletions in the low-coverage region. If the 1-bp indel region with the best alignments to the partial matches resulted in an average match of 35 out of 36 bp in all the sequence reads, then we considered this evidence of a real difference from the reference N2 sequence. For each of these 1-bp indels, we searched the unaligned reads from N2 for evidence of an identical polymorphism (again using an average match of 35 out of 36 bp as evidence for the polymorphism), because these 1-bp indels found in both LSJ2 and N2 sequencing are probable reference errors. We considered the remaining 41 1-bp indels as genuine differences between LSJ2 and N2 and classified them into the LSJ2 or N2 lineage using the unaligned MY14 sequencing reads as an outgroup.

The remaining low-coverage regions with partial matches were then visually inspected for the presence of larger deletions or insertions. The exact breakpoints for each deletion or insertion were defined using unambiguous regions, with MY14 as an outgroup, to classify the insertion or deletion into the N2 or LSJ2 lineage. A total of 26 indels larger than 1 bp were identified by this analysis.

A total of 331 indels were identified between LSJ1 and N2 in the previous whole-genome sequencing report[34]: a significantly higher number than the 67 indels identified here. Unlike the SNPs, we have not assessed the differences in indel predictions by PCR and Sanger sequencing.

**Deletion information.** Large deletions were verified using Sanger sequencing. The *srg-36/srg-37* region in LSJ2 contained a deletion of 4,906 bp replaced with

©2011 Macmillan Publishers Limited. All rights reserved

an AT nucleotide pair. The first 50 bp of the deleted region are gtgagagcggagtg atttcaaacacgggaatggccaaatacaacatcttc; the last 50 bp of the deleted region are tagt atgaacgattgaaaaaaatcaggccggctggaatcattggatatac. In CC1, 6,795 bp were deleted. The first 200 bp of the deleted region are gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtg tgtgtgtgtgtgt gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtg

relative to F0 was plotted individually for each trial. A second Matlab script was used to plot the average of all trials with standard errors for each time point.


31. Butcher, R. A., Fujita, M., Schroeder, F. C. & Clardy, J. Small-molecule pheromones that control dauer development in *Caenorhabditis elegans*. *Nature Chem. Biol.* **3,** 420–422 (2007).
32. Butcher, R. A., Ragains, J. R. & Clardy, J. An indole-containing dauer pheromone component with unusual dauer inhibitory activity at higher concentrations. *Org. Lett.* **11,** 3100–3103 (2009).
33. Li, H., Ruan, J. & Durbin, R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.* **18,** 1851–1858 (2008).
34. Weber, K. P. *et al.* Whole genome sequencing highlights genetic changes associated with laboratory domestication of *C.elegans*. *PLoS ONE* **5,** e13922 (2010).
35. Broman, K. W., Wu, H., Sen, S. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19,** 889–890 (2003).
36. Chalasani, S. H. *et al.* Dissecting a circuit for olfactory behaviour in *Caenorhabditis elegans*. *Nature* **450,** 63–70 (2007).


©2011 **Macmillan Publishers Limited. All rights reserved**